



<http://www.diva-portal.org>

Postprint

This is the accepted version of a chapter published in *HRI '17: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*.

Citation for the original published chapter:

Björnfot, P., Kaptelinin, V. (2017)

Probing the design space of a telepresence robot gesture arm with low fidelity prototypes

In: *HRI '17: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 352-360). ACM Digital Library

<https://doi.org/10.1145/2909824.3020223>

N.B. When citing this work, cite the original published chapter.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-139982>

# Probing the Design Space of a Telepresence Robot Gesture Arm with Low Fidelity Prototypes

Patrik Björnfot

Department of Informatics  
Umeå University  
901 87 Umeå, Sweden

patrik.bjornfot@umu.se

Victor Kaptelinin

Department of Informatics  
Umeå University  
901 87 Umeå, Sweden

victor.kaptelinin@umu.se

## ABSTRACT

The general problem addressed in this paper is supporting a more efficient communication between remote users, who control telepresence robots, and people in the local setting. The design of most telepresence robots does not allow them to perform gestures. Given the key role of pointing in human communication, exploring design solutions for providing telepresence robots with deictic gesturing capabilities is, arguably, a timely research issue for Human-Robot Interaction. To address this issue, we conducted an empirical study, in which a set of low fidelity prototypes, illustrating various designs of a robot's gesture arm, were assessed by the participants (N=18). The study employed a mixed-method approach, a combination of a controlled experiment, elicitation study, and design provocation. The evidence collected in the study reveals participants' assessment of the designs, used in the study, and provides insights into participants' attitudes and expectations regarding gestural communication with telepresence robots in general.

## Keywords

Telepresence robots; mobile remote presence; referential gestures; pointing; interaction design; low fidelity prototypes.

## 1. INTRODUCTION

Telepresence robots, also known as Mobile Remote Presence (MRP) systems, are a technology, which is increasingly common in many everyday settings, such as offices, classrooms, and healthcare environments [13] [16]. A telepresence robot is a device, which is remotely controlled by a person, often referred to as a "pilot", and serves as a physical avatar of the person, his or her embodied social proxy in a local setting.

A diversity of telepresence robots is currently available on the market [16]. The general outline of a telepresence robot is often similar to the general outline of a human body. A typical

telepresence robot comprises a wheeled base ("feet") and a camera/ display unit ("head"), connected to one another by a vertical pole.

Recently Mobile Remote Presence has become a topic of active research in Human-Robot Interaction (HRI) and Human-Computer Interaction (HCI) [2] [13] [26] [33]. In particular, studies have shown that telepresence robots can significantly improve the quality of interaction between a remote communication participant and local people. It was found that the use of telepresence robots supports ad hoc conversations (e.g., impromptu hallway discussions), makes the social presence of a remote person at the workplace more visible and appreciated by their colleagues (e.g., [31] [34]), and in other ways helps to compensate for negative aspects of working from a distance.

At the same time, the use of telepresence robots is associated with a number of problems, and further work is needed to find solutions to these problems. One of the problems, discussed in this paper, is a lack of gestural capabilities, typical of most telepresence robots. While, as mentioned, many robots look, very roughly, like humans, a noticeable difference is that robots typically do not have arms. There are some notable exceptions: for instance, a pioneering example of an experimental telepresence robot, named PRoP [24], was equipped with a pointing device resembling a human arm. However, the vast majority of popular telepresence robot models currently do not feature arm-like components.

The absence of arms means that robots cannot perform gestures, which is, arguably, an important limitation. Most telepresence robots are, in fact, *social* telepresence devices. Their primary, or often even sole, purpose is to support more flexible, physically situated, and embodied communication between remote pilots and local people. Given that gestures play a key role in interpersonal communication, it can be concluded that it is essential for robots to have gestural capabilities in order to fulfill their purpose.

The study reported in this paper aims to contribute to the development of gestural capabilities of telepresence robots by exploring the design space (or, rather, sub-space) of robot arms from an HCI perspective. The study focuses on "pointer arms", that is, arms, which are specifically intended to make it possible for the robot to perform deictic gestures.

The remainder of the paper is organized as follows. In the next section we present an overview of existing relevant research, clarify our intended contribution, and explain the rationale behind the choice of the methods used in our study. In the sections that follow we present, respectively, the methods used in the study, the results, and a discussion of our findings. We conclude with

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

HRI '17, March 06-09, 2017, Vienna, Austria

© 2017 ACM. ISBN 978-1-4503-4336-7/17/03...\$15.00

DOI: <http://dx.doi.org/10.1145/2909824.3020223>

general reflections on the study and an outline of issues for future research.

## 2. BACKGROUND

### 2.1 Telepresence robots as a communication technology: The importance of gestures

Telepresence robots in most everyday settings, such as offices, are almost exclusively used for communication. By making it possible for remote participants to actively choose their locations and perspectives, approach other people, e.g., to initiate an ad hoc conversation, and by increasing the visibility of a person in a setting, telepresence robots take a step beyond fixed-location videoconference systems. Arguably, a key direction of further development of telepresence robots should be making the robots a more effective and efficient communication technology. The aim of this paper is to explore one particular strategy of making telepresence robots a more powerful communication technology, namely, enhancing gestural capabilities of such robots.

Various strands of research in psychology, linguistics, ethology, and so forth, indicate that gestures play a fundamental role in human communication and cognition. For instance, Tomasello [32] argues that pointing and pantomiming were the first uniquely human forms of communication. McNeil [19] provides a thorough analysis of different types of gestures – iconic, deictic, beat, and metaphoric – and their integration with speech. A number of studies show that acquisition and mastering gestures is an essential aspect of human cognitive development, and that gestures not only express one's thoughts, but also can be considered a factor that influences the development of thinking itself [11]. However, as discussed below, so far the effort to support telepresence robots' gestural capabilities has been rather limited.

### 2.2 Related research: Gestures in HRI

Early in the development of telerobotic systems, both in academia and industry, there has been considerable interest in supporting gesture-based communication. A pioneering MRP system PRoP (Personal Roving Presence) featured an arm-like pointing device, which was considered an essential component of the overall design of the system [24]. Another relatively early system, PEBBLES [16] [33] had a rudimentary hand, which could be used for simple gestures, such as indicating a pupil's willingness to answer a teacher's question. The design of bidirectional communication robots CALLY and CALLO [36] employed a different approach, namely, using two identical robots in two different locations and making it possible for the user to control the remote robot by controlling the co-located robot, and vice-versa. An HRI-coaching system supporting full-body robot control is presented in [3]. Probably, the most advanced gesturing capabilities are implemented in the design of GestureMan, a special purpose telepresence robot (or rather several generations of robots), which is intended to support remote instruction on physical tasks [17].

Recently, however, a common approach has been to develop "armless" robots. Some robots, such as QB and MantaroBot [16] are equipped with laser pointers, but such devices only make it possible to highlight objects rather than perform pointing gestures, and the majority of systems do not feature any pointing devices at all.

At the same time, researchers in human-robot interaction and interaction design repeatedly emphasize the need for supporting gestural interaction. For instance, Tsui and Yanco [33] observe:

"As social interaction is the primary goal of social telepresence robots, failure to design for eye gaze, facial expressions, and nonverbal gestures will result in systems that hinder the ability to achieve telepresence for the user and/or the interactant".

The BEAMING project [7] emphasizes the need to support communicational and deictic gestures. Kaptelinin [15] argues that MRP systems should provide support for pointing, and outlines a set of requirements a pointing device for an MRP system should meet.

Therefore, enhanced gesture support can be considered a timely task in telepresence robotics research and development, a task, which is almost universally considered important but has not been sufficiently addressed so far.

There are two main streams of related research in gestures and robotics, firstly, the one related to how human gestures are interpreted by robots, e.g. [21][22] [30], and, secondly, the one related to gestures performed by robots when interacting with humans, e.g. [14]. This paper places itself in a third stream; we seek to find out how gestures can be mediated via robots in human-human interaction, with added challenges of making sure the gestures conform to social conventions.

The focus on performing deictic gestures (that is, pointing) via robot arms, which is somewhat similar to the focus of our study, characterizes some research dealing with assembly line arms, especially in scenarios where human intervention is needed [8], [9], [12], [29]. This research has produced some relevant findings: for instance, it was shown that gestures performed by robot arms should have good visibility and clarity. However, these analyses typically do not take into account social aspects of human-robot interaction, which aspects are central for our research.

### 2.3 Research methods: Experiment, gesture elicitation, and design provocation

A wide range of methods for user research, design, and evaluation have been developed in HCI and interaction design [27]. The present study employs low fidelity prototypes in combination with three empirical user research methods: controlled experiments, gesture elicitation, and design provocations. The choice of these methods is determined by the specific aims of our study and reflects the state of the art in HCI research methodology.

Low fidelity ("lo-fi") prototypes, as opposed to high fidelity ("hi-fi") prototypes, do not have the "look and feel" of the finished product. The value of sketches and lo-fi prototypes in design is often underestimated [5] and they can be considered "not serious enough". However, experience shows that lo-fi prototypes have a number of advantages, especially when engaging potential users early in the lifecycle of a design project. In particular, when working with lo-fi prototypes the participants are more inclined to suggest substantial changes to the designs they assess and focus on the underlying concept rather than the appearance of a prototype [27].

Controlled experiments, essentially modeled after the experimental method in psychology, have a long history in HCI [4]. Currently controlled experiments are, probably, less popular in HCI than studies conducted in natural setting, but they continue to play an important role in the field.

Gesture elicitation studies are intended to identify the most natural gestures, for different types of interaction between people and technology, by asking the participants to actually perform the gestures (e.g., [6]). As discussed below, the variant of the method used in our study was different from more traditional variants. Instead of asking people to show the gestures they themselves would perform by using *their own arms and hands*, we asked the participants to show what gestures they would expect a robot to perform, by physically adjusting the position of *the robot's arm*.

People who do not have prior experience with a novel technology can find it difficult to articulate a concrete opinion about the technology. Design provocations address this issue by presenting the participants with concrete designs, which designs can be (and often deliberately are) not optimal, to facilitate critical comments and suggestion improvements, and thus reveal participants' underlying assumptions and criteria for assessing the technology in question. Design provocations established themselves in HCI and interaction design as a part of wider user research and design methodologies, such as cultural probes [10], critical design [1][25], and reflective design [28].

The combination of methods used in the study was intended to provide different perspectives on the same object of analysis. This approach follows the general idea of “method triangulation”, which is considered an important strategy for increasing the external validity of a study [23].

### 3. METHOD



Figure 1. Experimental setting.

### 3.1 Participants

Eighteen participants, from 21 to 30 years of age (average age: 25 y. o.), 10 males and 8 females, took part of the study. Most of the participants were students at a Swedish university.

### 3.2 Materials

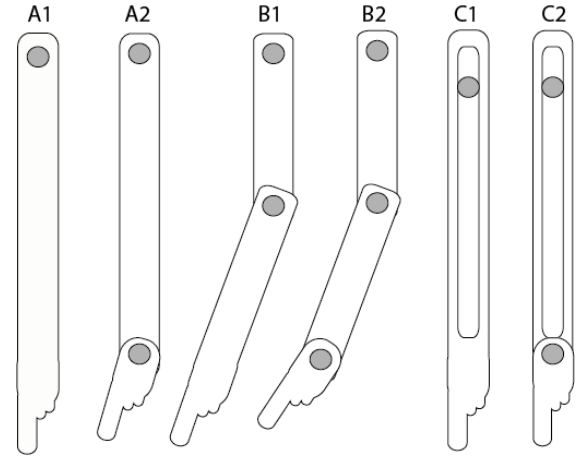


Figure 2. The arm designs used in the study.

#### 3.2.1 Prototypes

The experiment employed low fidelity prototypes (Figure 1) implementing six different designs of a robot arm (Figure 2). The designs were selected because they are characterized by low complexity and therefore represent a natural starting point for exploring robot arms' gestural expressivity. The designs were produced by combining two factors: (a) the construction of the arm (“Arm Type”), and (b) whether or not the “hand” was a fixed part of the arm or could be moved independently (“Independent Hand”).

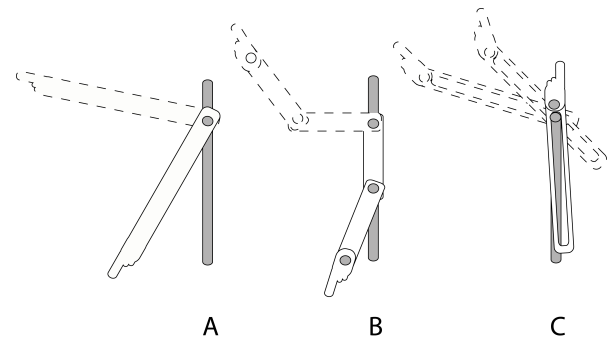
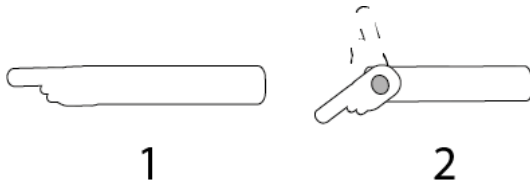


Figure 3. Arm Types: Fixed Attachment Stick (A), Elbow Joint (B), and Sliding Attachment Stick (C).

There were three levels of the Arm Type factor, which corresponded to the following variants of arm construction: a one-piece arm with a fixed connection to the robot body (“Fixed Attachment Stick”, Figure 3A), a two-piece arm having an equivalent of an elbow joint in the middle (“Elbow Joint”, Figure 3B), and a one-piece stick with a sliding connection to robot's body (“Sliding Attachment Stick”, Figure 3C). The Independent Hand factor had two levels, Yes or No (see Figure 4).



**Figure 4. Independent Hand Factor levels: No Independent Hand (1), Independent Hand (2).**

In the discussion below the six prototypes are coded using letters and numbers. The letters, A-C, signify three different arm types: Fixed Attachment Stick (A), Elbow Joint (B), and Sliding Attachment Stick (C). The numbers, 1 and 2, signify two levels of the Independent Hand factor: absence (1) or presence (2) of an independent hand. For instance, C1 means “Sliding Attachment Stick without an independent hand”.

All six low fidelity prototypes were implemented using identical wireframe stands, 67 cm high, and identical “head parts” made of polystyrene foam. The arms were made of durable glossy cardboard; wing nuts were used to conveniently adjust the connection between different parts of an arm (Figure 1).

### 3.2.2 Imaginary use scenarios

Three different use scenarios of a telepresence robot performing a pointing gesture were presented to the participants. The first scenario (“Can I talk to you?”) described a robot approaching a group of people and indicating an intention to talk to one person from the group. The second scenario (“Look at this!”) described a robot referring to an object in the local environment, a gadget lying on a desk, during a conversation with a local person. The third one (“Mind the door”) described a robot approaching a door and indicating the intention to “touchlessly” open it (as if using a “remote control”). The scenarios did not specify the pilot’s user interface.

The choice of the scenarios was informed by empirical and conceptual analyses of telepresence robots [16][33], which identify a number of issues, common for various use contexts. Three such issues, reflected in the experimental scenarios, were: (a) addressing a specific person, (b) referring to an object in the local setting, and (c) dealing with obstacles while navigating a telepresence unit.

## 3.3 Experiment design

The experiment employed a two-factor within-subject experimental design. Six experimental factors corresponded to the six robot arm designs described in section 3.2.1; they were produced by combining three levels of the Arm Type factor and two levels of the Independent Hand factor. To minimize potential order effects, a Latin Square plan was used to determine the sequence of conditions for individual participants.

## 3.4 Procedure

Each of the participants was tested individually by a team of experimenters, each team comprising 3-4 persons. The experimenters were master’s level students taking a course in user research; running an experimental study was a course assignment. The experimenters received a training on how to conduct the study; the training included one of the researchers going through a complete trial experimental session with each of the experimenter teams.

Each experimental session comprised three parts, corresponding to the three scenarios described in section 3.2.2. In each of these parts the participants were first required to read a scenario description. The descriptions were provided as printed documents to avoid potential researchers’ biases. Then the participants were asked to explore the six prototypes (by following a predefined order), fill in a form, adjust the position of each of the prototype arms to show what the participant thought was the most natural gesture for the prototype, and take part in a brief interview. The total time for each individual session was about 40 min.

## 3.5 Types of empirical data collected in the study

### 3.5.1 Assessment with numerical scores

For each of the three imaginary use scenarios the participants were asked to provide two different types of scores. First, they were supposed to assign ranks to the six designs, ranging from “1”, which indicated the best design overall, to “6”, indicating the worst design. Second, after demonstrating the most appropriate gestures for each of the six prototypes, the participants were asked to use four seven-point Likert scales to assess the gestures according to the criteria of clarity, confusability, politeness, and perceived safety. The scales’ values were ranging from “-3” to “+3”, with more positive scores corresponding to more positive assessments. The selection of the criteria was informed by a previous study by Kaptelinin [16].

### 3.5.2 Elicited gestures

Each of the participants was required to physically manipulate the prototypes to show what gestures they thought were the most appropriate ones for each of the three scenarios. In total, the participants produced 324 gestures, which were photographed by the experimenters.

### 3.5.3 Interviews

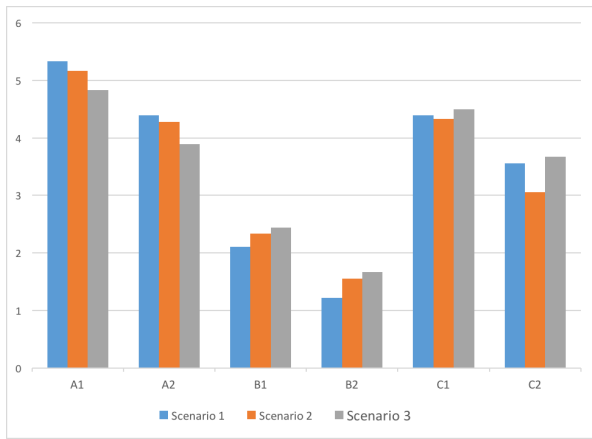
In addition to providing numerical scores and producing gestures, the participants were required to explain the reasons for their ranking, assessments, and the choice of gestures. The participants were also encouraged to express their opinions about the designs, suggest improvements, and share their general reflections about gesture-based interaction with telepresence robots.

## 4. RESULTS

### 4.1 Numerical scores

#### 4.1.1 Overall ranking

Figure 5 provides an overview of how the participants ranked the six prototypes used in the study, separately for each of the three scenarios. The figure shows that the most highly rated design was the most human-like one, B2 (“Elbow Joint with an independent hand”). Seven participants ranked B2 the highest in all three scenarios. The lowest ranks were given to A1 (“Fixed Attachment Stick without an independent hand”). The ranks of other prototypes were distributed between these two extremes.



**Figure 5. Mean ranks of the prototypes, separately for each of the three scenarios (shorter bars correspond to higher ranks).**

Three patterns in the overall ranking results deserve special attention. First, the order of prototypes, from best to worst, was the same in all three scenarios: B2, B1, C2, A2, C1, A1. Second, within all arm types (A, B, and C) the prototype with an independent hand received a higher average rank than the prototype without an independent hand. Third, the results suggest that there are individual differences in the ranking: while B2 was ranked as one of the top two (#1 or #2) most often (48 times), other prototypes also received top ranks: B1 (41 times), C2 (10 times), C1 (6 times), A1 (3 times), and A2 (1 time).

#### 4.1.2 Assessment of visibility, clarity, politeness, and perceived safety

The results of the assessment of the prototypes, in each of the scenarios, according to their visibility clarity, politeness, and perceived safety, are consistent with the data from overall ranking. The Arm Type B (“Elbow Joint”) received highest mean ranks, and was followed by C and A, according to both overall ranking and scale-based assessment (see Table 1). The mean ranking is consistent across all scenarios.

**Table 1. Mean four-scale scores and mean ranks for three Arm Types.**

Model	Visi- bility	Clari- ty	Polite- ness	Feels Safe	Total Value	Rank
A	2,2	1,6	0,3	0,6	4,7	4,6
B	2,4	2,2	1,7	1,8	8,0	1,9
C	2,0	1,6	0,6	0,8	5,0	3,9

The arms with independent hands were rated higher than arms without independent hands, again according to both overall ranking and scale-based assessment (see Table 2). The ranking is consistent across all scenarios.

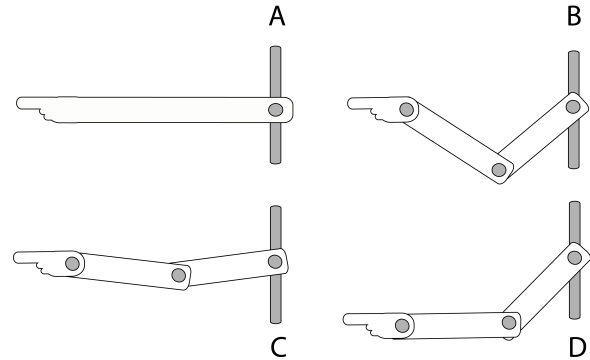
**Table 2. Mean four-scale scores and mean ranks for two levels of the Independent Hand factor.**

Independent- hand	Visi- bility	Clari- ty	Polite- ness	Feels Safe	Total Value	Rank
No	2,2	1,7	0,7	1,0	5,5	3,9
Yes	2,2	1,9	1,0	1,1	6,3	3,0

The total sum value of a condition for all four criteria correlates with the mean rank of the condition.

## 4.2 Elicited gestures

### 4.2.1 Scenario 1 (“Can I talk to you?”)



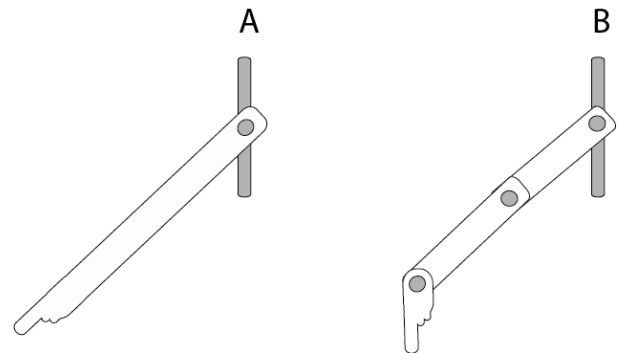
**Figure 6. Gestures produced by the participants in Scenario 1.**

The most common gesture elicited in Scenario 1 was a straightforward pointing (see Figure 6A). It was suggested in 25% of the cases. Similar to this gesture was a slightly downwards pointing gesture, suggested in 8% of the cases. These two gestures have a mean ranking of, respectively, 4,4 and 4,7 (where 1 is best, and 6 is worst).

When producing gestures with Arm Type B (“Elbow Joint”) prototypes, 77% of the participants used the shoulder joint to move the arm down, while the “forearm” and hand were still pointing forward. These gestures generally had a better ranking. Eight percent of the gestures included the “upper arm” directed down at 45°, the “forearm” raising up at 45°, and the hand pointing straight ahead (Figure 6B). In 6.5% of the cases the whole arm basically pointed straight ahead, with just a slight bent of the elbow (Figure 6C). In additional 6% of the gestures the upper arm was directed down at 45°, while the forearm and hand were pointing straight ahead (Figure 6D).

Seven of the participants commented that it was impolite to point at someone, and three of these participants did not even want to point. The participants thought that it was more polite to point with a non-straight arm, and all of these seven participants used the elbow joint on B-prototypes to bend the arm.

### 4.2.2 Scenario 2 (“Look at this!”)



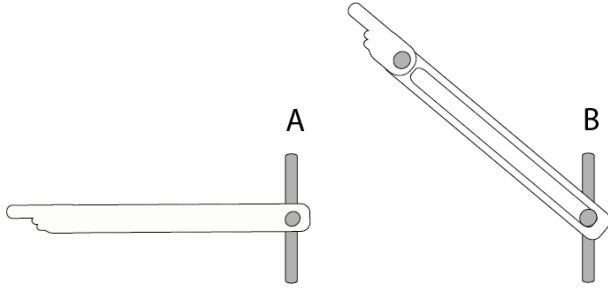
**Figure 7. Gestures in Scenario 2.**

Gestures elicited in Scenario 2 were rather homogeneous. In 56.4% of the cases the proposed gesture was the one shown in Figure 7A: the arm and hand were pointing down at 45°. In 9% of

the cases the arm was pointing at 45° down, and the hand was pointing straight down (Figure 7B). The mean rankings of both of these gestures were close to average (3,72 and 2,9, while the average was 3).

Half of the participants indicated in their comments that a key factor when performing the ranking of the prototypes was the flexibility of the arms (5 additional participants mentioned it in relation to other scenarios). For instance, one of the participants, who had all prototypes do exactly the same gesture, stated that they rated the arms only on the basis of their flexibility, and therefore rated the B2 the highest, followed by B1.

#### 4.2.3 Scenario 3 (“Mind the door”)



**Figure 8. Gestures in Scenario 3.**

In this scenario 38% of the gestures were simply pointing straight ahead (the mean ranking was 5,17), see Figure 8A. In 14% cases the arm was pointing up at 45 degrees (the mean ranking was 3), see Figure 8B.

A concern raised by three of the participants was that it was hard to make a gesture that signals “the door is about to open” and not “I need help to open the door”.

### 4.3 Suggestions for improvement

In addition to ranking, scale-based assessment, and gesture elicitation the participants, in each of the three scenarios, were also encouraged to critically evaluate the proposed designs and suggest possible ways of improving them. The suggestions for improvement are summarized below and listed in Table 1.

In Scenario 1 (“Can I talk to you?”) the majority of suggestions were related to politeness. The suggested improvements were: (a) providing a more complex hand, which would comprise more than just an index finger, (b) making it possible for the robot to rotate the hand about the arm’s axis, (c) provide fully flexible ball bearing joints (1 participant), and (d) support pointing with face recognition.

By proposing an increased flexibility of the arm the participants tried to avoid what was considered a “rude pointing”. One participant noted that “It is difficult to be polite when you point with a finger”, and suggested that this issue can be addressed by supporting either the use of an open hand or subtle movements of the pointing finger.

The reason for suggesting face recognition functionality was to simplify pilot’s task of pointing to a particular person.

Most of the suggestions in Scenario 2 (“Look at this!”) were dealing with clarity, visibility, and solving the problem of precision. The suggested improvements were: (a) using a LED/Laser pointer, (b) making the shoulder height adjustable, (c) holding a remote in robot’s hand, (d) including lights that would visually highlight the arm, (e) when pointing, using the entire

hand instead of a finger. The LED/Laser pointer was supposed to increase the accuracy of pointing, and adjustable shoulder height was meant to increase the flexibility of pointing to meet the requirements of different tasks and environments. To point with a full hand instead of a finger was the only suggestion that reduced the precision of pointing.

Improvement suggestions in Scenario 3 (“Mind the door”) mostly focused on visibility. They included: (a) horizontal (e.g., left to right) movement of the arm, e.g., to signal that the local user should move out of the way, (b) rotation of the hand about the arm’s axis, and (c) using robot’s “head screen” for displaying animation. In addition, two participants highlighted the need for supporting repetitive movements.

As a general suggestion, two participants proposed using LED lights on the robot as “warning signals” of different types of actions.

**Table 3. Summary of improvement suggestions.**

Name	# of Participants	Scenario
Providing a more advanced hand, not just a finger	6	1 & 3
A two DoF hand (combining vertical and horizontal movements)	5	1 & 3
Fully flexible, three DoF ball joints	1	1
Supporting pointing with face recognition	1	1
Employing a LED/Laser pointer	3	2
Making the shoulder joint height adjustable	1	2
Placing a remote control in robot’s hand	1	2
Pointing with a hand instead of finger	1	2
Lighting up the arm	1	2
Horizontal arm movement	4	3
Using robot’s screen for animations displaying intended actions	1	3
Using LED light to signal events	2	Other

### 4.4 Other comments of the participants

#### 4.4.1 Human likeness

In their general comments and reflections, most participants (14 out of 18) expressed their positive assessment of human likeness of telepresence robots. As one participant observed: “The more human the better”. This notion is supported by the fact that Arm Type B (“Elbow Joint”) prototypes, which were described by the participants as “human-like”, were also ranked and assessed most positively. In a similar vein, one participant suggested that Arm Type C prototypes (“Sliding Attachment Stick”) should have a shoulder in order to “look more human”.

#### 4.4.2 Participants role in the scenarios

It is evident that even though the participants were instructed to take the role of the local user, they provided comments and reflections from the pilot’s side, as well. Most notable was the suggestion to use face-recognition for a more efficient detection of potential communication partners.

## 5. DISCUSSION

### 5.1 Design provocation

The scenarios used in the study made the participants question the proper way to interact via an MRP system. Participants' comments highlighted the importance of cultural norms (e.g., "It is impolite to point to a person" or "I would not point at people at all"), and, at the same time, the participants demonstrated and explained how to increase politeness (e.g., by using an open hand, avoiding direct pointing, waving, etc.). Arm Type A ("Fixed Attachment Stick") designs in Scenario 1 served as designed provocations: they made the participants realize that a straight arm can be considered "authoritarian", "commanding" and "rude". However, even though the expressive capabilities of Arm Type A designs were mostly considered essentially inadequate, the participants had no choice but use them to produce "the best possible" gestures because it was required by the instructions. Included in the ranking, the gestures provide a baseline for comparison for gestures produced with other designs.

When encountering problems caused by the limitations of the arm designs used in the study, the participants suggested a number of improvements. Several of such improvements were in line with making robot arms more human-like. However, there were also other types of suggested improvements, which had no references to the human body. For example, different types of lights and lasers, as well as various ways of utilizing the robot's "head" screen, were proposed. These types of improvements, which did not imply human likeness, would not be proposed if MRP systems were assumed to have the same capabilities as human beings. It appears that the suggestions, instead, reflect an assumption that regarding nonverbal communication capabilities telepresence robots may have specific advantages to be exploited, not limited to just replicating human capabilities.

### 5.2 Flexibility and human-likeness

Flexibility was a key theme in participants' comments related to Scenario 2, even in cases when this capability was not actually used. The most plausible explanation appears to be that the participants were thinking about similar scenarios, in which extended flexibility would have been needed. One could argue that the rating forced the participant to go beyond the context of the scenario at hand in order to make a more general comparison between the designs to establish the "winner". This explanation is currently just a hypothesis, which needs to be investigated empirically. It should be mentioned that several of the participants provided improvement suggestions that intend to increase arm's flexibility, such as a rotating hand, horizontal movement, or adjustable shoulder height.

Another desirable attribute of telepresence robots that was commonly mentioned by the participants was "human-likeness", mainly in relation to politeness, but also in relation to forming gestures in general.

Judging from participants' comments, the results can be interpreted as an indication that human-likeness and flexibility were considered desirable qualities, which positively affected the rating regardless of whether the capabilities were actually used in gesture elicitation. The arm type that was most human-like, B2, had more joints than other arm types, and was therefore also the most flexible one. An interesting issue to explore would be a comparison of B2 with other designs, featuring more joints (and thus being less human-like), in order to separate human-likeness from flexibility.

### 5.3 Is human-likeness optimal?

It is natural that the human body provides a basic point of reference when looking at social communication via MRP systems. After all, as humans, we are mostly used to socially interacting with other humans.

The comment "the more human the better", expressed by some participants, is an important indication; it shows that the participants think that achieving human-likeness is a key direction of further improvement of telepresence robots. However, striving for human-likeness faces two significant challenges: first, the need to explain the phenomenon of Uncanny Valley and, second, deal with the issue of "false expectations".

The Uncanny Valley phenomenon means that if a robot resembling a human still lacks some human traits, it can elicit negative feelings in people who perceive it [20]. It also means that when striving for human likeness robot designers run a very real risk of making people feel uncomfortable.

The other factor that needs to be taken into account when considering human-likeness as a design objective is that technology, in this case, the "arm", may imply more human capabilities than it actually has, which may give rise to false expectations. For example, the suggestion that the arm should feature a "full hand", if implemented, can make local people interacting with a telepresence robot think that the arm has the capability of picking up things, even if the arm can only be used for gestures.

In light of the above arguments, it can be concluded that the role and effect of human-likeness in mobile remote presence needs to be studied more thoroughly, for instance by exploring robot arm designs, which have identical functional capabilities but differ in "human likeness".

### 5.4 Gestures and culture

Our study suggests that people use their cultural norms and expectations, such as those related to "politeness", when imagining and modeling robot arm gestures. However, the study does not present any concluding evidence regarding the specific relationship between culture and gesture-based communication. It was not an object of analysis in our study, and further research is needed to explore this issue.

In particular, the cultural background of the participants was rather homogeneous; the majority of them were students at a Swedish university. In future studies involving participants with more diverse cultural backgrounds can be a way to better understand the role of culture in mobile remote presence.

An important issue for further research is also the dynamics of cultural expectations and norms over time. Studies have shown that such expectations and norms develop when telepresence robots are introduced to real-life settings and activities [18], and one can expect similar phenomena to take place when/if gesture-based communication with telepresence robots becomes a common aspect of everyday contexts. Emerging social practices are likely to be reflected in new or modified norms, e.g., regarding what is considered polite.

## 6. CONCLUSION

The underlying idea of the study reported in this paper is that designing gesture arms for telepresence robots is an important issue for Human-Robot Interaction, and addressing this issue requires a systematic exploration of the entire design space of

such gesture arms. Some findings of our study, e.g., that the participants preferred a human-like arm with an elbow joint and that designs with an independent hand systematically showed an advantage (if a modest one) over designs without such hand, can be of certain interest to the designers of robot arms. Admittedly, the study explores only a small subset of relevant design issues, and further work is needed to provide a more substantial guidance for design. However, the intended contribution of our study extends beyond an assessment of concrete design solutions.

The evidence collected in the study allows us to make a general assessment of the research strategy we adopted, that is, the use of low fidelity prototypes. The evidence suggests that low fidelity prototypes – which can be produced quickly and are inexpensive – can facilitate participants' engagement and constructive critical assessment of design solutions. The unfinished nature of the prototypes stimulates imagination, and physical embodiments of design concepts support the participants in making specific critical comments and suggestions for improvement.

At the same time, our evidence suggests that low-fidelity prototypes have some substantial limitations. In particular, while the prototypes were effective in bringing in the issue of politeness, the issue of safety, which is arguably no less important, was not a prominent theme in the discussion. A possible explanation is that the small form factor of the prototypes did not make it possible for the participants to *experience* telepresence robots as a source of potential danger.

In sum, our study suggests that design explorations of telepresence robot arms, based on low-fidelity prototypes, can be a useful strategy. Using such prototypes as a first step when exploring a novel design space is a fast and cheap way to define the direction of research and design. It makes it possible to freely and efficiently, if tentatively, explore a wide range of issues, identify some of the key ones, facilitate participants' feedback, and stimulate the search for new alternatives. At the same time, studies employing low fidelity prototypes cannot be expected to provide definitive answers to research and design questions. They need to be complemented with more advanced prototypes and products, which would make it possible to extend the scope of analysis to include more developed technological solutions and more realistic user experiences, actions, and contexts.

## 7. ACKNOWLEDGMENTS

We want to thank Umeå University students taking the “User Research for Interaction Design” course (Fall 2016) for running the experiments and four anonymous reviewers for their valuable comments. The work reported in this paper was funded by The Swedish Research Council (Vetenskapsrådet), grant 2015-05316.

## 8. REFERENCES

- [1] Bardzell, J., and Bardzell, S. 2013. What is “critical” about critical design? *Proceedings of CHI '13*. ACM New York, 3297-3306 DOI= <http://dx.doi.org/10.1145/2470654.2466451>
- [2] Beer, J. M. and Takayama, L. 2011. Mobile remote presence systems for older adults: Acceptance, benefits, and concerns. *Proceedings of HRI 2011*. NJ: IEEE. DOI= <http://dx.doi.org/10.1145/1957656.1957665>
- [3] Bogdanovych, A., Stanton, C., Wang, X., and Williams, M.-A. 2012. Real-time human-robot interactive coaching system with full-body control interface. In *Robot Soccer World Cup XV*. Lecture Notes in Artificial Intelligence. Springer, Berlin, 562-573.
- [4] Blandford, A., Cox, A. L., and Cairns, P. A. 2008. Controlled Experiments. In Cairns, P.A., & Cox, A.L. (eds.). *Research Methods for Human Computer Interaction*. CUP.
- [5] Buxton, B. 2007. *Sketching User Experience: Getting the Design Right and the Right Design*. Morgan Kaufmann, SF.
- [6] Chan, E., Seyed, T., Stuerzlinger, W., Yang, X.-D., and Maurer, F. 2016. User Elicitation on Single-hand Microgestures. *Proceedings CHI 2016*. DOI= <http://dx.doi.org/10.1145/2858036.2858589>
- [7] Cohen, B., Lanir, J., Stone, R., and Gurevich, P. 2011. Requirements and design considerations for a fully immersive robotic telepresence system. In *Proceedings of HRI 2011 Workshop on Social Robotic Telepresence* (pp. 16-22).
- [8] Ende, T., Haddadin, S., Parusel, S., Wüsthoff, T., Hassenzahl, M., and Albu-Schäffer, A. (2011, September). A human-centered approach to robot gesture based communication within collaborative working processes. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3367-3374). IEEE. DOI= <http://dx.doi.org/10.1109/IROS.2011.6094592>
- [9] Gleeson, B., MacLean, K., Haddadi, A., Croft, E., and Alcazar, J. (2013, March). Gestures for industry: intuitive human-robot communication from human observation. In *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction* (pp. 349-356). IEEE Press.
- [10] Gaver, B., Dunne, T., and Pacenti, E. (1999). Cultural probes. *interactions* 6 (1), Jan./Feb., 21– 29.
- [11] Goldin-Meadow, S. 2015. Gesture and cognitive development. In *Handbook of Child Psychology and Developmental Science*. V. 2. Wiley.
- [12] Haddadi, A., Croft, E. A., Gleeson, B. T., MacLean, K., and Alcazar, J. (2013, May). Analysis of task-based gestures in human-robot interaction. In *Robotics and Automation (ICRA), 2013 IEEE International Conference on* (pp. 2146-2152). IEEE. DOI= <http://dx.doi.org/10.1109/ICRA.2013.6630865>
- [13] Herring, S.C., Fussell, S. R., Kristoffersson, A., Mutlu, B., Neustaedter, C., and Tsui, K. 2016. The Future of Robotic Telepresence: Visions, Opportunities and Challenges. *Proceeding of CHI EA '16*. NY: ACM Press, 1038-1042. DOI= <http://dx.doi.org/10.1145/2851581.2886423>
- [14] Huang, C.-M., and Mutlu, B. "Modeling and Evaluating Narrative Gestures for Humanlike Robots." *Robotics: Science and Systems*. 2013.
- [15] Kaptelinin, V. 2016. Supporting Referential Gestures in Mobile Remote Presence: A Preliminary Exploration. In *Inclusive Smart Cities and Digital Health*. LNCS, v. 9677, Springer International Publishing, 262-267. DOI= [http://dx.doi.org/10.1007/978-3-319-39601-9\\_23](http://dx.doi.org/10.1007/978-3-319-39601-9_23).
- [16] Kristoffersson, A., Coradeschi, S., and Loutfi, A. 2013. A review of mobile robotic telepresence. *Advances in Human-Computer Interaction* 2013, 3, nnn-nnn.
- [17] Kuzuoka, H., Oyama, S., Yamazaki, K., Suzuki, K., and Mitsuishi, M. 2000. GestureMan: A Mobile Robot that

- Embodies a Remote Instructor's Actions. *Proceedings CSCW'00*, December 2-6, 2000, Philadelphia, PA. 155-162. DOI= <http://dx.doi.org/10.1145/358916.358986>
- [18] Lee, M. K. and Takayama, L. 2011. "Now, I have a body": uses and social norms for mobile remote presence in the workplace. *Proceedings CHI 2011*. DOI= <http://dx.doi.org/10.1145/1978942.1978950>
- [19] McNeill, D.D. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.
- [20] Mori, M., MacDorman, K. F., and Kageki, N. (2012). The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100.
- [21] Nehaniv, C. L., Dautenhahn, K., Kubacki, J., Haegele, M., Parlitz, C., and Alami, R. A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction." *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. IEEE, 2005.
- [22] Obaid, M., Kistler, F., Häring, M., Bühling, R., and André, E. (2014). A framework for user-defined body gestures to control a humanoid robot. *International Journal of Social Robotics*, 6(3), 383-396. Doi > 10.1007/s12369-014-0233-3
- [23] Oulasvirta, A. 2009. Field experiments in HCI: Promises and challenges. In P. Saariluoma, H. Isomaki (Eds.), *Future Interaction Design II*. Springer. DOI= 10.1007/978-1-84800-385-9\_5
- [24] Paulos, E. and Canny, J. 1998. PRoP: Personal Roving Presence. *Proceedings CHI 98*. DOI= <http://dx.doi.org/10.1145/259081.25919>
- [25] Pierce, J., Sengers, S., Hirsch, T., Jenkins, T., Gaver, W., and DiSalvo, C. 2015. Expanding and Refining Design and Criticality in HCI. *Proceedings of CHI 15*. 20183-2092. DOI= <http://dx.doi.org/10.1145/2702123.2702438>
- [26] Rae, I., Mitlu, B., and Takayama, L. Bodies in motion: Mobility, presence, and task awareness in telepresence. *Proceedings of CHI'14*. ACM Press, NY (2014). DOI= <http://dx.doi.org/10.1145/2556288.2557047>
- [27] Rogers, Y., Sharp, H., and Preece, J. 2015. *Interaction Design: Beyond Human-Computer Interaction*. 4<sup>th</sup> edition. Wiley.
- [28] Sengers, P., Boehner K., David S., and Kaye J. 2005. Reflective design. *Proceeding CC '05 Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*. ACM. New York, 49-58.
- [29] Sheikhholeslami, S., Moon, J., and Croft, E.A. 2015. "Exploring the effect of robot hand configurations in directional gestures for human-robot interaction." *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015. DOI= <http://dx.doi.org/10.1109/IROS.2015.7353879>
- [30] Stanton, Christopher, Anton Bogdanovych, and Edward Ratanasena. "Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning." *Proc. Australasian Conference on Robotics and Automation*. 2012.
- [31] Takayama, L. and Go, J. (2012). Mixing metaphors in mobile remote presence. *Proceedings of CSCW '12*. NY: ACM Press. DOI= <http://dx.doi.org/10.1145/2145204.2145281>
- [32] Tomasello, M. *Origins of Human Communication*. MIT Press, Cambridge, Mass. (2008)
- [33] Tsui, K.K. M. and Yanco, H. A. 2013. Design challenges and guidelines for social interaction using mobile telepresence robots. *Reviews of Human Factors and Ergonomics* 9, 1, 227-301. DOI= <http://dx.doi.org/10.1109/ROMAN.2004.1374827>
- [34] Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G. G., Lee, B., and Inkpen, K. 2010. Embodied social proxy: Mediating interpersonal connection in hub-and-satellite teams. *Proceedings of CHI '10*. NY: ACM Press, 1049-1058. DOI= <http://dx.doi.org/10.1145/1753326.1753482>
- [35] Yim, J.-D., and Shaw, C. D. (2011). Design considerations of expressive bidirectional telepresence robots. *Proceeding of CHI 2011, Extended Abstracts*. NY: ACM Press, 781-790. DOI= <http://dx.doi.org/10.1145/1979742.197963>