UNIVERSITAT POLITÈCNICA DE CATALUNYA

Doctoral Programme:
AUTOMATIC CONTROL, ROBOTICS AND COMPUTER VISION

Ph.D. Thesis

DESIGNING A ROBOT TO EVALUATE GROUP FORMATIONS

Dennys Paillacho Chiluiza

Advisor:
Prof. Cecilio Angulo Bahón
Co-advisor:
Prof. Marta Díaz Boladeras

March, 2019

Abstract

Robots are making their way in environments inhabited by people. Whether in domestic or public crowded environments, robots should take into consideration social norms and behaviors in order to become a social robot. This dissertation focuses on the problem of how to develop a robotic platform in order to validate human-robot interaction experiences in realistic environments. More specifically, we are concerned with social interactions in human-robot groups in public scenarios, where a variety of people can converge. Our final goal is the develop of a social robot based on certain theories of group behavior and the use of space, known as spatial relationships. The intermediate goals are related with the design and development of the experiences in the wild: as minor changes as possible in the scenario, definition of social tasks, gradual development of a robotic platform as transparent as possible from the robotic side.

Initially, this research introduces several preliminary studies of human-robot interaction (HRI) with the PAL Robotics' REEM robot at the CosmoCaixa Science Museum in Barcelona. Based on certain theories about the use of space as a form of social communication or interaction, the task under study with the commercial robot is as a museum guide, both when the group was in motion (*i.e.* when it was being guided) as well as a group in a static place. Moreover, a second HRI study with REEM robot accomplishing the task of a teacher's assistant was carried out to analyze the perception of the robot's social presence and identity.

Likewise, the development of a robotic platform, known as MASHI, for the study of HRI is presented. Based on the service to be completed by the robot, improvements in the experimental robotic platform (structure, morphology, head, face, arms) were carried out in continuous cycles following

the development of HRI experiences. This structure should be hold as simple as possible in order to make it ‘transparent’ in the social HRI study.

Next, the field study of human-robot social interaction with the MASHI robot with the role of exhibition guide in a cultural center is presented. Based on direct observation techniques, a study is made of the different spatial relationships that are generated when a robot interacts with a person or groups of people.

Finally, a novel approach to represent the spatial relationships of sHRI in a qualitative way is introduced for future experiences. In this concluding study, we analyze different spatial arrangements generated in a social scenario with a robot within the guide role.

As a main conclusion, it can be stated that people follow social norms, in the form of spatial relationships, when interacting with a robot that provide a social service in a public space. Children, however, recurrently challenge these social norms, probably because they are constantly learning about the norms that regulate our coexistence.

Spatial relationships are clearly reinforced when the role assigned to the robot is more explicit and understood by people. Spatial relationships can be affected by the characteristics of the environment, either by the available space or by the elements arranged in it, as well as by the number of people who inhabit it.

Overall, this dissertation points out that the provided service, and its understanding from the user’s side, is more important that the robotic skills of the robotic platform in order to improve user experiences in public environments.

Resum

Els robots s'estan fent en entorns habitats per persones. Ja sigui en entorns domèstics o públics, els robots haurien de tenir en compte les normes i els comportaments socials per convertir-se en un robot social. Aquesta tesi es centra en el problema de com desenvolupar una plataforma robòtica per validar experiències d'interacció humà-robot en entorns realistes. Més específicament, ens preocupa les interaccions socials en grups humans-robots en escenaris públics, on una varietat de persones poden convergir. El nostre objectiu final és desenvolupar un robot social basat en determinades teories del comportament grupal i l'ús de l'espai, conegudes com a relacions espacials. Els objectius intermedis es relacionen amb el disseny i el desenvolupament de les experiències a la natura: com canvis menors en l'escenari, definició de tasques socials, desenvolupament gradual d'una plataforma robòtica tan transparent com sigui possible des de la part del robot.

Inicialment, aquesta investigació introdueix diversos estudis preliminars d'interacció humans-robot (HRI) amb el robot REEM de PAL Robotics al CosmoCaixa Science Museum de Barcelona. Partint de determinades teories sobre l'ús de l'espai com a forma de comunicació social o interacció, la tasca que s'estudia amb el robot comercial és una guia del museu, tant quan el grup estava en moviment (*i.e.* quan també es guiava) com a grup en un lloc estàtic. A més, es va dur a terme un segon estudi HRI amb el robot REEM que va realitzar la tasca d'un assistent de professor per analitzar la percepció de la presència i identitat social del robot.

Així mateix, es presenta el desenvolupament d'una plataforma robòtica, coneguda com MASHI, per a l'estudi de l'HRI. Basant-se en el servei que ha de completar el robot, es van realitzar millores en la plataforma robòtica experimental (estructura, morfologia, cap, cara, armes) en cicles continus

seguint el desenvolupament d'experiències HRI. Aquesta estructura hauria de ser tan simple com sigui possible per fer-la 'transparent' en l'estudi HRI social.

A continuació, es presenta l'estudi de camp de la interacció social entre humans i robots amb el robot MASHI com a guia d'exposició en un centre cultural. Basant-se en tècniques d'observació directa, es fa un estudi de les diferents relacions espacials que es generen quan un robot interactua amb una persona o grups de persones.

Finalment, s'introdueix un enfocament innovador per representar les relacions espacials del sHRI d'una manera qualitativa per a les experiències futures. En aquest estudi final, s'analitzen els diferents arranjaments espacials generats en un escenari social amb un robot dins del rol de guia.

Com a conclusió principal, es pot afirmar que les persones segueixen normes socials, en forma de relacions espacials, quan interactuen amb un robot que proporciona un servei social en un espai públic. Els nens, tanmateix, desafien de forma recurrent aquestes normes socials, probablement perquè estan constantment aprenent sobre les normes que regulen la nostra convivència.

Les relacions espacials es veuen clarament reforçades quan el paper assignat al robot és més explícit i entès per les persones. Les relacions espacials poden veure's afectades per les característiques del medi ambient, ja sigui per l'espai disponible o pels elements disposats en ell, així com per la quantitat de persones que hi habiten.

En general, aquesta tesi assenyala que el servei proporcionat i la seva comprensió per part de l'usuari és més important que les habilitats robòtiques de la plataforma robòtica per millorar les experiències dels usuaris en entorns públics.

Resumen

Los robots se abren paso en entornos habitados por personas. Ya sea en entornos domésticos o públicos, los robots deben tener en cuenta ciertas normas y comportamientos sociales para convertirse en un robot social. Esta disertación se centra en el problema de cómo desarrollar una plataforma robótica para validar experiencias de interacción humano-robot en entornos realistas. Más específicamente, nos preocupamos por las interacciones sociales en grupos humano-robot en escenarios públicos, donde una gran variedad de personas puede converger. Nuestro objetivo final es el desarrollo de un robot social basado en ciertas teorías de comportamiento grupal y el uso del espacio, conocidas como relaciones espaciales. Los objetivos intermedios están relacionados con el diseño y desarrollo de las experiencias 'en la naturaleza': cambios mínimos como sea posible en el escenario, definición de tareas sociales, desarrollo gradual de una plataforma robótica lo más transparente posible desde el lado robótico.

Inicialmente, esta investigación presenta varios estudios preliminares de interacción humano-robot (HRI) con el robot REEM de PAL Robotics en el Museo de Ciencias CosmoCaixa de Barcelona. Basado en ciertas teorías sobre el uso del espacio como una forma de comunicación o interacción social, la tarea en este estudio con el robot comercial es como guía de museo, tanto cuando el grupo estaba en movimiento (es decir, cuando estaba siendo guiado) como cuando el grupo estaba en un lugar estático. Además, se llevó a cabo un segundo estudio de HRI con un robot REEM que realizaba la tarea de un asistente de profesor para analizar la percepción de la presencia e identidad social del robot.

Asimismo, se presenta el desarrollo de una plataforma robótica, conocida como MASHI, para el estudio de la HRI. En función del servicio que debe

completar el robot, las mejoras en la plataforma robótica experimental (estructura, morfología, cabeza, cara, brazos) se llevaron a cabo en ciclos continuos siguiendo el desarrollo de las experiencias de HRI. Esta estructura debe mantenerse lo más simple posible para que sea 'transparente' en el estudio de HRI social.

A continuación, se presenta el estudio de campo de la interacción social humano-robot con el robot MASHI con el papel de guía de exposición en un centro cultural. Con base en técnicas de observación directa, se realiza un estudio de las diferentes relaciones espaciales que se generan cuando un robot interactúa con una persona o grupos de personas.

Finalmente, se introduce un enfoque novedoso para representar las relaciones espaciales de la sHRI de forma cualitativa para las experiencias futuras. En este estudio final, analizamos diferentes arreglos espaciales generados en un escenario social con un robot con el rol de guía.

Como conclusión principal, se puede afirmar que las personas siguen normas sociales, en forma de relaciones espaciales, cuando interactúan con un robot que brinda un servicio social en un espacio público. Los niños, sin embargo, desafían recurrentemente estas normas sociales, probablemente porque están aprendiendo constantemente sobre las normas que regulan nuestra convivencia.

Las relaciones espaciales se refuerzan claramente cuando el rol asignado al robot es más explícito y entendido por las personas. Las relaciones espaciales pueden verse afectadas por las características del entorno, ya sea por el espacio disponible o por los elementos dispuestos en él, así como por el número de personas que lo habitan.

En general, esta disertación señala que el servicio prestado, y su comprensión del lado del usuario, es más importante que las habilidades robóticas de la plataforma robótica con el fin de mejorar las experiencias del usuario en entornos públicos.

To my dear parents Elvia and Nilo.

Acknowledgements

My thanks to all those people who have somehow stepped up their shoulders so that this thesis work has gone forward.

To my thesis directors, Cecilio Angulo and Marta Díaz, for giving me the opening to undertake this line of research; his advice and guidance allowed me to explore new paths of knowledge.

To Knowledge Engineering Research Group - GREC, for giving me the opportunity to present my work and provide feedback on it.

Jorge Álvarez and Carlos Gómez for hosting and undertaking the development of the MASHI robotic platform.

To the students of the UPC of different subjects that made their contribution in the development of MASHI, such as Javier Poveda, Mathilde Hamici, Pablo Pardo, Alex Pardo, Marc Bolaños, Lukáš Danev, Marten Hamann, Nicolas Fricke, Tobias Hollarek, among others.

A great appreciation to L'Hospitalet City Hall in Barcelona, Spain for providing the research facilities used in this study. I'm particularly grateful for the encouraging support of Mr. Ricardo Castro at La Bòbila Cultural Center.

A special thanks to all those institutions that with their financing have made possible the accomplishment of this thesis work, such as the Escuela Superior Politécnica del Litoral - ESPOL, Fundación Carolina, Secretaría de Educación Superior, Ciencia, Tecnología e Innovación - SENESCYT of Ecuador as well as the PATRICIA Research Project (TIN2012-38416-C03-01,03), funded by the Spanish Ministry of Economy and Competitiveness.

Finally, and no less important, an infinite thank you to Rosa Almachi, my traveling companion, and to all my family, thank you for your continued support.

List of Figures

List of Tables

LIST OF TABLES

Chapter 1

Introduction

1.1 Scope and Motivation

In our days, a growing number of robots are being used not only in industrial environments, but also in environments where people live, such as in the domestic environment or in public spaces. As robotic applications have moved closer to the activities of people's daily lives, robots should be able to communicate and interact with people in a close and fluid way [1], referred in the literature to as "*Social Robots*". In this sense, Human-Robot Interaction - HRI includes several factors such as the social, emotional and cognitive aspects of the interaction. Areas such as rehabilitation, independent living, marketing, education, customer service, security, museum guide, leisure and entertainment, among others, are potential fields of application for social robots.

In some of these areas of application the issue of mobility in a social robot is important. Key questions to be addressed for an effective performance of social robots featured with this capability are: how to move (*i.e.* speed, kind of movement, trajectories), where to perform (*i.e.* proximity management) and how to place (*i.e.* distance, position, pose and orientation) to be unobtrusive and socially congruent. Promising attempts to optimize social robots spatial management in different scenarios (*e.g.* assistive telepresence at home) have been developed applying models and knowledge from social psychology (*i.e.* proxemics, space formations, group walking patterns and crowd dynamics)[2, 3, 4]. Socially compliant navigation [5] implies planning and performing robot's trajectories and motion behavior taking into account the communicative function and social rules of space management in a shared location. Smart spatial behavior (*e.g.* , interpersonal distance, orientation) according to social norms would not only enhance collocated user's safety and acceptance but also provide mobile robots with nonverbal channels to communicate intentions (*e.g.* shift direction, initiate displacement) and to express emotional content [6]. This thesis study human-robot interaction in real and social environments from the perspective of the use of the space as a form of social interaction.

In order to have a simplification of the outside world, social robots was most often evaluated within labs and in constrained environments [7]. The study of sHRI in natural environments, on the other side, presents several challenges, among which we can mention: (a) the variety and complexity of the environment; (b) the multimodality of the interaction; (c) the diversity of robots; and (d) the variety of people. Hence, due to the dimension of the diversity of scenarios, the solutions proposed in the literature are constrained to specific domains. For instance, from the environment point of view, proposed solutions are intended to work in public, but not too crowded spaces. Regarding the different modalities in interaction, they include studies on verbal and non-verbal interaction (e.g. facial expressions, proxemics, spatial formations). In addition, different robotics platforms have been considered in the social human-robot interaction framework, which can be classified according to: their shape, such as human-like robots and pet-like robots; their size, such as human-sized robots and desk-sized robots. Finally, there is a wide range of users that should be considered in this problem. Different solutions have been presented for young and elder people because it would depend on their profile for a successful implementation. Another factor to take into account is the number of participants in the interaction.

The main motivation to carry out this research work is the need to carry out HRI experiences in real environments and where there is a variety of people, since this will allow us to have a more comprehensive vision of the environment where the interaction takes place, on the user (or groups of users) that interact with the robot and about the design of the robot itself. Based on this premise, we study the design of a robot centered on the human and based mainly on the interaction experiences that are being carried out, which will allow us to know which are the main characteristics to take into account in the design and construction of a social robot.

1.2 Objectives

The main objective of this dissertation is the design of a social robot that offers different types of services in open environments where it is expected that a group of people interact with the robot, and that can serve as a basis for the study of human-robot spatial interactions.

The specific objectives of this thesis are:

1. To explore social human-robot interactions in a natural setting with a standard commercial robot;

2. To design and construct a robotic platform for the study of social human-robot interactions based on the robot service;

3. To carry out studies of human-robot interactions with the developed robotic platform in real environments; and,

4. To explore an approach to represent spatial human-robot interaction from a qualitative perspective.

1.3 Thesis Outline

The thesis is organized as follows. Chapter 2 reviews the state-of-the-art of human-robot interactions in open environments. Several studies of human-robot interaction in a naturalistic setting, the Science Museum of Barcelona, is developed in Chapter 3. In Chapter 4, the design and construction of a robotic platform, called MASHI, used for the study of social human-robot interactions is presented. In Chapter 5, several studies of social human-robot interaction with the MASHI robotic platform in a Cultural Center are presented were carried out. Chapter 6 discusses a novel approach to represent social Human-Robot Interaction spatial relationships using Qualitative Spatial Representation theory. Finally in Chapter 7 conclusions of this thesis, as well as future works are presented.

1.4 Main Outcomes

1.4.1 HRI research.

Several exploratory studies of human-robot interaction in real environments have been carried out, which help to enrich the knowledge on how to evaluate the social interactions in these scenarios. In addition, as a result of these experiences, a database for future HRI studies has been collected.

Furthermore, based on the theory of qualitative spatial representations, a novel model to represent the spatial relationships that occur during social interaction in groups of people-robot was proposed.

1.4.2 HRI design.

During the iterative studies, improvements were also made in the definition of robot services, as well as in the establishment of scripts and flow diagrams to carry out social interaction.

1.4.3 Robot design.

During this research a multiproposite robotic platform for the study of human-robot social interaction was developed. From the experiences carried out, development cycles were carried out to improve or add more functionalities to the platform. This is how a light teleoperated robotic platform with anthropomorphic proportions was obtained with a modular design and with the ability to interact through verbal and non-verbal language.

1.4.4 Publications.

During the period of research in the Automatic Control Department, at the Universitat Politècnica de Catalunya, the author has been published several parts of this thesis in journals, international conferences, and technical reports. A list of these works is given below:

- **HRI Conf 2014.** Díaz-Boladeras, M.; Paillacho, D.; Angulo, C.; Torres, O.; González, J.; Albo-Canals, J. "A Week-long Study on Robot-visitors Spatial Relationships During Guidance in a Sciences Museum". Proceedings of the 2014

ACM/IEEE International Conference on Human-robot Interaction, ACM, 2014, 152-153

- **IJHR Journal 2015.** Díaz-Boladeras, M.; Paillacho, D.; Angulo, C.; Torres, O.; Gonzalez, J.; Albo-Canals, J. "Evaluating Group-Robot Interaction in Crowded Public Spaces: A Week-Long Exploratory Study in the Wild with a Humanoid Robot Guiding Visitors Through a Science Museum". International Journal of Humanoid Robotics, 2015, 12, 1550022

- **IROS Conf 2015.** Paillacho, D.; Angulo, C.; Díaz-Boladeras, M. "An exploratory study of group-robot social interactions in a cultural center". IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS 2015 Workshop on Designing and Evaluating Social Robots for Public Settings, 2015, 44-48

- **JARCA Conf 2016.** Paillacho, D.; Falomir, Z.; Angulo, C. "Towards modelling group-robot interactions using a qualitative spatial representation". Proceedings of the XVIII Workshop on Qualitative Systems and Applications in Diagnosis, Robotics and Ambient Intelligence (JARCA 2016), CEUR Workshop Proceedings (CEUR-WS.org), 2016

- **ETCM Conf 2017.** Danev, L.; Hamann, M.; Fricke, N.; Hollarek, T.; Paillacho, D. "Development of animated facial expressions to express emotions in a robot: RobotIcon". 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM), 2017, 1-6

Chapter 2

State of the Art

As service robots are increasingly closer to the activities of daily living, making their way towards the so-called social robot, they must have the ability to interact with people closely and fluently [1] both in verbal and non-verbal way; which raises the importance and the need to study social Human-Robot interaction in real environments.

In this sense, Section 2.1 several concepts around human-robot interaction in real environments are reviewed. Next, in Section 2.2 a review of the service robots used in interaction and guidance tasks in open and real environments is performed, with special emphasis on the methods and techniques used to evaluate the social interaction. Finally, Section 2.4 describes the theoretical foundations to evaluate social Human-Robot interaction through its spatial relationships; as well as the HRI studies conducted in this regard.

2.1 HRI Concepts

2.1.1 Social robot

According to the different modes of interaction, different classes of social robots are proposed in [8], going from the most basic to the most complex social scenarios:

- Socially evocative: designed to encourage people to anthropomorphize the technology to interact with it.
- Social interface: robots that use social signals similar to human ones and communication modes to facilitate interactions with people.

- Socially receptive: more perceptive robots to human social signals, allowing people to manipulate robot behavior in different ways.

- Sociable: robots that actively participate with people in order to, not only benefit the person, but also to benefit itself. These robots not only perceive human social signals, but at a deeper level, they generate social models of people to interact with them.

2.1.2 Wizard of Oz - WoZ

In real environments, in a dynamic physical scenario and with a multitude of people, it is currently very complex to operate a "sociable" autonomous robot. For this reason, HRI studies often include a human as a part of the system, through the Wizard of Oz - WoZ technique [9, 10], in which a human operator, behind the curtains, controls the robot without the knowledge of the users.

2.1.3 Open Vs Closed environments

As mentioned in [11], in real situations social robots may deploy their activity in close proximity to people either in closed (e.g. home [12], school [13], office, nursing home [14]) and open (e.g. exhibitions [15], museums [11, 16, 17], malls) environments. In closed environments the occupants are known, and often belong to a few homogeneous profiles (e.g. ages, familiarity to technology).

On the other hand, in open public spaces occupants are unknown, diverse, variable and dynamic, often including heterogeneous profiles (i.e. teenagers, staff, elderly). A frequent situation in large-scale open public environments is the configuration of dense crowds that the robot is supposed to travel through fulfilling safety (the primary requirement of a robot operating in a public space), reliability and social requirements at a time. In this sense, there were two situations that have been distinguished in human-robot groups: fixed groups (mentioned as "interaction" in [11]), when the space occupied by the Human-Robot group normally does not change in the time (as in the case of a robot performer-human audience); and mobile group (mentioned as "guidance" in [11]), that is, when the space occupied by the group changes continuously over time (as in the case of a robot guiding people).

2.1.4 HRI types

Regarding the different approaches to the study of HRI, [18] mentions three types:

Robot-centered.- It emphasizes the point of view of a robot as a creature, i.e. as an autonomous entity that pursues its own objectives based on its own motivation, actions and emotions.

Human-centered.- It is primarily concerned with how the robot can perform its task in an acceptable and comfortable way to the person. In this category, studies point to how people react and interpret the robot's appearance and behavior, regardless of the behavioral architecture of the robot or the cognitive processes that may occur within the robot.

Centered on robot cognition: In this case the robot is emphasized as an intelligent system, that is, a machine that makes decisions for itself and that solves problems and faces them as a part of the tasks to be performed.

2.1.5 Social situation

Based on the explanatory models of interpersonal relationships in [19], it is offered in [1] an interesting frame of reference for understanding interaction and designing it, starting from the description of the *social situation* in which the robot is going to behave. The social situation, according to these models, is defined by the interrelation between the agents (eg, child-therapist), the role assumed by the robot, the scenario - both physical and social (eg, accompanied only in the hospital, at home) - and the purpose of the activity (choosing a product, performing rehabilitation exercises).

2.1.6 Social presence

In the study of the HRI, one of the concepts that is handled with higher frequency is the one of "social presence". The concept of presence originally refers to two different phenomena. The first one is related to the sensation of being present in a virtual environment. It can be defined as "the feeling of being there" [20]. The second one can be related to the feeling of being in the company of a social entity: the perceptual illusion of non-mediation [21]. In this context, the second definition will be used along this dissertation.

2.1.7 Robot identity

Naming a robot reflects and creates its identity, as well as illustrates the social relationships between the robot designer and the robot [22]. Coeklebergh [23] argues that human-robot relations are mediated by language and that our conversation towards or about robots is not a mere representation, nor only a participation in a social act, but it reshapes our relationship with "others". Analyzing the linguistic construction in terms of the robot's identity (such as name, age and sex) of the people involved in a human-robot social interaction can serve as a means to evaluate the subject's degree as a social/human entity.

2.2 Guiding Service Robots

Guidance is one of the most useful services robots may deliver in public spaces as museums, exhibitions, malls, and tourist sites. Taking the role of guide, the robot not only provides people with appropriate information to make the visit a more enjoyable experience, but help them to get to intended destinations. The main difference between a promoter versus a guidance robot is that guidance in public spaces also implies social navigation in a highly dynamic scenario. In this section some tour guide robots are described making special emphasis on the user interface as a means for human-robot interaction.

In 1997, Burgard *et al.* presents the museum guide robot RHINO, tested at the Deustche Museum in Bonn during six days. In [24], Burgard *et al.* describe the software architecture of RHINO, presenting the estimation modules of planning and execution states and the human-robot interaction. The robot has two user interfaces, one on-board and one web interface. Among the points of discussion is the interactive component of the robot, where it is indicated that user interaction is essential for the overall validity of the concept, emphasizing the need to develop friendly and intuitive interfaces as key prerequisites if the robots are to be part of the daily life of people.

In 1998, Thrun *et al.* presents a second generation of RHINO, called MINERVA, that was put into operation for two weeks at the Smithsonian's National Museum of American History in Washington. The paper describes the software architecture considering aspects of navigation in dynamic as well as unmodified environments, short-term interaction and virtual telepresence [25]. Thrun points out that interaction with people is the main objective of MINERVA, and that the type of interaction was characterized by the following factors: museum visitors do not have any knowledge of robots and that they can not be instructed previously on its use, that the robot must interact with both a person and a group of people and that most people spend less than 15 minutes with the robot, what they call "short-term interaction". Among the improvements in MINERVA, from the point of view of interaction, are the development of a face, direction of the head and the use of the voice. In addition, a stochastic finite state automaton of 4 states was used to model simple emotional states, which depending on the block on the robot's route ranged from happy to angry. Among the results, of a total of 63 people consulted, 36.9% believed that MINERVA's level of intelligence is comparable to that of a human. Likewise, 27% of all people (mainly children under 10 years of age) thought that MINERVA was alive, while 69.8% thought it was not. Finally, the importance of having a voice recognition system to improve the interactive capabilities of the robot is mentioned.

In 2002 Siegwart *et al.* presents the guiding robot ROBOX, that was tested at the Swiss National Exhibition Expo02 during several months. The goal of the ROBOX was to allow the visitor a unique experience in sharing the environment with intelligent mobile robots and to interact with them. In this sense, the authors use the Scenario Object Utility Language (SOUL), in order to implement different scenarios using all the input and output channels available in the robot. In this experience, there was a several issues as: several people prefer to play and interact with the robot more than been guided at the exhibition, some visitors were also not very patient and not willing to follow the robot, due to the people crowding at the exhibition, the robot voice was difficult to understand, plenty of visitors need to get information about the robot technology; and some visitors has high expectations about the robot functionality.

In [17], Willeke *et al.* presents an evolutionary study of a series of museum robots: CHIPS (1998), SWEETLIPS (1999), JOE HISTORYBOT (1999) and ADAM 40-80 (2000). The robots have been presented in a variety of venues, as the Carnegie Museum of Natural History, the Heinz History Center, the Republican National Convention, the Democratic National Convention, a shopping mall, the National Aviary and the Pittsburgh International Airport. [17] presents a series of lessons learned as a result of HRI experiences as: in public space, there were often a crowd of people around the robot, rather than a single person; with the background noise of the environment, this makes more difficult to people to understand the robot; long presentations drive the audience away; a same response dialogue seems more scripted and less spontaneous and finally the psychological effect of creating an anthropomorphic robot.

In [11], Shiomi *et al.* describe a field trial with the interactive humanoid robots ROBOVIE and ROBOVIE-M at the Osaka Science Museum in 2007. The visitors used RFID tags, whose information was used to plan the robot interaction with the visitors. In the interaction, the robots use gestures and utterances resembling the free play of children. Additionally, the robots performed exhibition guiding by moving around several exhibits and explaining it based on sensor information. An experiment from the field trial revealed that the combination between free-play interaction and exhibit-guiding positively affects the visitor experience at the museum. The work mention several social abilities in the robots for open environments as: during a robot explanation visitors were distracted trying to evoke reactions from the robot; name-calling behavior from the RFID tag was very distracting in crowds of people; robot's exercise behavior and hugging behavior entertaining has a good impact in children, but not in adult-only groups.

In [26], Karreman *et al.* presents a user evaluation study with the autonomous tour guide robot FROG at the Royal Alcázar in Seville in 2014. Several innovations in navigation and vision were integrated in the design of the robot behavior to provide interactive tours and adaptive content to visitors. The study mentions that even if isolated technical features work perfectly in controlled settings, they might not work well in the real world situation, because the unexpected behaviors of people. This is the case of the facial expression recognition system, which detected the facial expressions when the visitor was right in front of the robot. in a real situation, where groups of people were around the robot, a greater distance between the robot and the visitors was generated, so the system did not work. The authors recommend implementing in-the-wild studies from early stages in the development process, in such a way that the user's behavior can be understood in order to create a robot capable of handling this situation.

As it has been observed, the habitual environments of the museum guide robots can be very dynamic, with a multitude of people and of different ages. In this case, conducting an evaluation of the HRI in this type of environment constitutes a challenge. Below are some methods and techniques used to evaluate HRI.

2.3 Metrics and Methods to Evaluate HRI

In the design of a social robot not only the fulfillment of the task assigned to the robot must be considered, but also that the interaction carried out between the robot and the person must be fluid [27] [28]. For this, it is necessary to be able to evaluate the HRI in some way. Depending on the object of the study, to evaluate the HRI, it is necessary to establish what is to be measured (*i.e.* the metrics) and how it is to be measured (*i.e.* the method).

In this sense, Olsen and Goodrich propose some metrics to evaluate the quality of the human-robot interface, such as: task effectiveness, neglect tolerance, robot attention demand, free time, fan-out and interaction effort [29]. These metrics are intended to measure the effort of interaction on the part of the user. In this study, the interaction is limited to the fulfillment of a certain task, and the robot is considered as a servant who simply receives and executes orders of the person. Applied to our guide robot model, the defined metrics would be more focused on the interaction between the operator and the robot; also that in the case of robot-user interaction the problem would be how to apply these metrics to the group-robot interaction, since these metrics take into consideration only the interaction between a robot and a person.

Steinfeld *et al.* propose a set of metrics based on the tasks carried out by mobile robots [30]. Steinfeld *et al.* proposes five types of tasks: navigation, perception, management, manipulation and social. From the point of view of our model, both the navigation and the perception will be performed by a human operator, the robot does not perform any type of manipulation, and the administration of the robot will be carried out by a single operator. There will be no manipulation tasks by the robot. In the social part, Steinfeld raises the difficulty of determining the metrics (engineering, psychological, sociological) more appropriate to assess social effectiveness in an HRI. Among them, the author mentions several metrics as: characteristics of the interaction, persuasiveness, truthfulness, engagement and compliance.

In [28], Brayda and Chellali mention two types of metrics: overt and covert. Among the overt measures are gestures, gesticulations, facial expressions, inter-person distance, verbal speech, eye gaze, body language, and impulsive movements. Covert measures are more difficult to infer, such as affective states, emotional reactions (trust, sense of safety) and non-verbal messages, as well as automatic responses such as body temperature, skin pressure, and other biological signals that include the electrical activity of the brain.

Weiss *et al.* propose a framework for evaluating human-robot collaboration, which is composed of four evaluation factors (abb.USUS): usability, social acceptance, user experience and societal impact [31]. These evaluation factors are focused on the user, so they can be suitable for our study. The term usability refers to the ease of using an object, and this concept has different metrics, such as: effectiveness, efficiency, learnability, flexibility, robustness and utility. The concept of social acceptance in this context is defined as "an individual's willingness based on interaction experiences to integrate a robot into an everyday social environment". Among the metrics to evaluate social acceptance are: performance expectancy, effort expectancy, attitude towards using technology, self efficacy, forms of grouping, attachment and reciprocity. Among the user experience metrics are: embodiment, emotion, human-oriented perception, feeling of security, and co-experience with robots. Finally, the social impact factor has the following metrics: quality of life, health and security, education and cultural context.

Establishing the importance of standardizing HRI measurement tools, Bartneck *et al.* proposes five user-centered metrics to evaluate HRI: anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety [32]. For this, Bartneck *et al.* proposes the use of questionnaires based on semantic differential scales.

From a more integral perspective, Young *et al.* indicates that human-robot interactions are unique personal experiences, due to the complexity of the context [33]. In this sense, Young *et al.* introduces a set of three perspectives to study social interaction with robots: visceral factors of interaction (*e.g.* immediate response), social mechanics (*e.g.* the application of social norms) and the social macro-level structures. These perspectives can be used by expert evaluators from either the human or the robot point of view, based on user experience or robot design respectively. To evaluate the HRI from this perspective, we usually use qualitative techniques such as the thick description, collect multiple points of view or more structured approaches such as Grounded Theory, culture or technology probes or contextual design.

In [7], Sabanovic *et al.* mention the importance of carrying out HRIs in open environments, outside the laboratory. For the evaluation of HRI in real environments, Sabanovic *et al.* proposes different types of metrics, such as: interaction spaces, group interaction, interpersonal interaction, and rhythmicity; the same ones that use the systematic observation method. This approach is commonly used in psychology, ethology, and sociology to study observable behavior such as activities, postures, gestures, facial expressions, movements, and social or human system interactions. It should be performed by social scientists trained in observational analysis.

In the line of overt measurements, the use of space as a means to evaluate HRI is raised in [34]. In this paper, a series of metrics based on human spatial behavior are discussed. In particular, the individual, attentional, interpersonal and physiological factors that contribute to the social space are analyzed.

In an attempt to measure the interaction quality of a mobile telepresence robot, in a semi-structured environment, Kristoffersson *et al.* correlates the spatial formations (seen later in the next section) established between a telepresence robot and a user (actor) and the perceived presence and ease of use of the operator [2]. In this study, systematic observation is used for the analysis of spatial formations; as well as two questionnaires (Temple Presence Inventory and Networked Minds Social Presence Inventory) to evaluate the perceived presence and ease of use of the operator.

A compendium of the different user-centered/whole HRI metrics (an extension of the USUS framework presented in [31]) as well as the methods that can be used to evaluate the HRI is shown in Table 2.1. According to our robot-group interaction model in open environments, we are interested in knowing the perceptions of the users, but also the behavior of the robot-group. In this sense, the most suitable methods would be the use of questionnaires and systematic observations. However, in open and crowded environments, conducting a robot-group survey would be a challenge.

In the next section we detail an approach to evaluate HRI through the spatial relationships.

2.4 Spatial Relationships in HRI

Social robots as physical entities that co-inhabit a place with people in HRI (eventually, sHRI) are involved in what is known as *spatial relationships* [2, 12]. Spatial relationships, a mode of non-verbal communication, are a combination of distance, relative position and spatial arrangement that occur naturally whenever two or more people engage in

Table 2.1: User-centered/whole HRI metrics and methods used to evaluate HRI

	Methods							
Research Objectives/Metrics	**Expert Eval**	**User Studies**	**Quest.**	**Physio. Measures**	**Focus Groups**	**Interviews**	**Conversat. Analysis**	**Systematic Observation**
Social task [30]								
Characteristics of the interaction							x	x
Persuasiveness							x	x
Truthfulness			x					
Engagement								x
Compliance								x
Usability [31]								
Effectiveness	x	x						
Efficiency	x	x						
Learnability	x	x						
Flexibility	x	x						
Robustness	x	x						
Utility			x			x		
Social Acceptance [31]								
Performance Expectancy			x		x			
Effort Expectancy			x		x			
Self Efficacy			x		x			
Forms of Grouping			x		x			
Attachment			x		x			
Reciprocity			x					
User Experience [31]								
Embodiment			x		x			
Emotion			x	x	x			
Human-Oriented Perception			x					
Feeling of Security			x	x	x			
Co-Experience			x		x			
Societal Impact [31]								
Quality of Life			x		x	x		
Working Conditions			x		x	x		
Education			x		x	x		
Cultural Context			x		x	x		
Key concepts in HRI [32]								
Anthropomorphism			x					
Animacy			x					
Likeability			x					
Perceived intelligence			x					
Perceived safety			x					
Holistic view of HRI [33]								
Visceral factors of interaction	x							
Social mechanics	x							
Social macro-level structures	x							
HRI in real environments [7]								
Interaction spaces								x
Group interaction								x
Interpersonal interaction								x
Rhythmicity								x
Human spatial behavior [34]								
Individual								x
Attentional								x
Interpersonal								x
Physiological								x
Quality of a mobile telepresence robot [2]								
Spatial formations								x
Perceived presence			x					
Ease of use of the operator			x					

an interaction [35] and convey significant and relevant social information (e.g. how each of them is involved) and also define an interpersonal space for developing activity.

An open public scenario where autonomous mobile robots have been deployed are museums. Three aspects make the robot navigation in a museum specially difficult: the robot has to guide visitors through dense even crowded spaces, some elements of the physical space could be “invisible” to the robot (e.g. glass walls) and the configuration of the environment change frequently (e.g. pieces of furniture, fences). The robot guide in a museum faces two primary challenges: navigating safely, reliably and socially through crowds, and interact with people in a compelling and intuitive way [16].

Service robots featured with *walk around* functionality must deal with crucial social navigation issues: how to move (i.e. speed, kind of movement and trajectories), where to perform and how to place (i.e. distance, position and orientation) to be unobtrusive, effective and socially congruent. Moreover, robots in shared spaces get involved in *spatial relationships* with people [2, 12]. Spatial relationships are a combination of distance, relative position and orientation that occur naturally whenever two or more people engage in an interaction [35] and convey significant and relevant social information and also define an interpersonal space for developing activity [2]. Promising attempts to optimize social robots spatial management in different scenarios (*e.g.* telepresence assistance at home) have been done applying models and knowledge from social psychology (*i.e.* proxemics, space formations, group walking patterns and crowd dynamics).

Many disciplines can contribute to the understanding of spatial relationships in HRI in open and real scenarios. Below relevant concepts such as proxemic behavior, F-formations and group behavior are introduced and discussed.

2.4.1 Proxemics

The term *proxemics* was introduced by anthropologist Edward T. Hall in 1966 [36] to refer to *“the interrelated observations and theories of man’s use of space as a specialized elaboration of culture”* [ibid, p. 1]. In this regard, Hall defines 4 kinds of interpersonal distances, each with its own significance in a social context: intimate ($0 - 0.45$ meters), personal ($0.45 - 1.2$ meters), social ($1.2 - 3.6$ meters) and public (> 3.6 meters)(see Figure 2.1). These interpersonal distances may vary depending on culture.

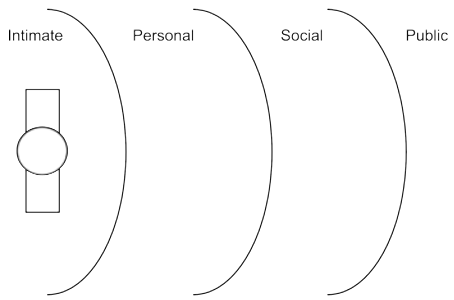


Figure 2.1: Proxemic distances.

2.4.2 F-Formations

The *F-formation* system was proposed by Adam Kendon [37] to mention the spatial arrangement, both in position and orientation, that are generated when two or more people interact in a face to face way and affirm that *"behaviour of any sort occurs in a three dimensional world and any activity whatever requires space of some sort "* [ibid, p. 1.] This space allows an organism to perform any activity and is differentiated from other spaces [35]. According to Kendon, in any scenario is common that several individuals are co-present, but the way they are positioned and oriented in relation to the others reflects directly how they can be involved together. Based on his observations, Kendon defines a transactional space, known as *o-space*, defined as the space where people can interact and manipulate shared objects. Kendon described different types of arrangements or spatial formations (see Figure 2.2):

- 'vis-a-vis'.- individuals who are facing one each other. This formation can only be seen in dyadic groups.
- 'L-shape' - individuals that are oriented perpendicularly to each other. This formation can only be seen in dyadic groups.
- 'Side-by-side'.- individuals that are oriented parallel to each other and facing to the same direction.
- 'Circular form'.- when the individuals are positioned around to and facing the *o-space*. Typically in groups of three or more individuals.
- 'Horseshoe shape'.- a semicircular arrangement, a kind of compromise between side-by-side and circular form.

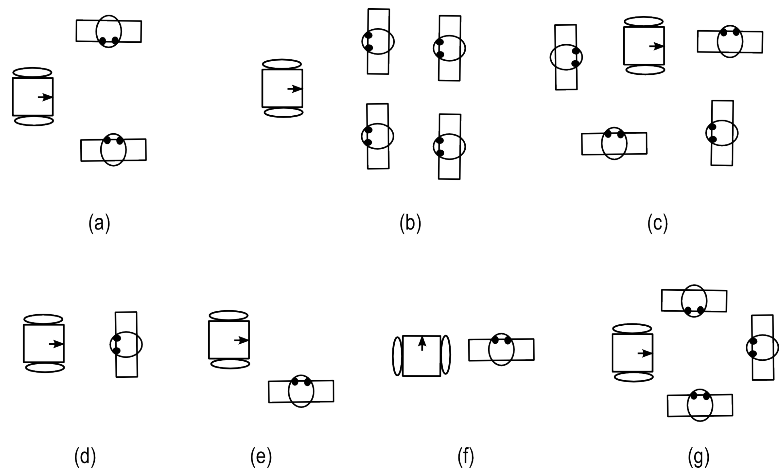


Figure 2.2: Face-to-face spatial arrangements (F-Formations): (a) 'Horseshoe shape'; (b) 'Performer-audience'; (c) 'Cluster'; (d) 'vis-a-vis'; (e) 'L-shape'; (f) 'Side-by-side'; (g) 'Circular form'.

- 'Performer-audience'.- spatial arrangement where there is an unequal distribution of rights to start a conversation or action. Typically in larger groups.
- 'Cluster' .- refer to a group of people that do not follow any spatial arrangement.

Empirical studies in robotic applications have studied the spatial relationships of a human-robot group as a main issue in order to improve the quality of interaction, taking into account that interpersonal distances convey significant and relevant social information [2]. An interesting conclusion is that when physical constraints (e.g. narrow passages) in combination of navigational requirements unable the robot to maintain the convenient spatial behavior, it can compensate this situation with other interactive behaviors (e.g. verbally apologizing for an inappropriate distance or reducing the eye-contact) to maintain an overall degree of desired intimacy.

Although there are several studies that evaluate the HRI using the spatial relationships, all are limited to experimental environments only in the individual and in closed and in experimental environments [2, 4, 12, 38, 39, 40].

2.4.3 Walking Behavior

This section addresses the identification and description of space relationships during guidance applying models from social psychology such as group walking [41] and crowd dynamics [42].

Guidance is a demanding collaborative task that requires communicating intentions [43] (i.e. robot offers the service, visitors select a destination and request the "Bring me there" function, the robot heads towards the destination) and social navigation (i.e. walk together to the target location). Walking along following the leader implies complex space regulations (i.e. distancing, spatial configurations) to allow guide and visitors group up and walk together effectively. These space relationships during guidance must be at a time socially meaningful and compatible with the robot's navigation specifications (i.e. collision avoidance performance [44]).

To model the navigation through crowds of dynamic agents with uncertain trajectories some attempts has been done drawing inspiration from the pedestrians behaviors in dense environments, where people usually engage in "joint collision avoidance" (called the social forces model) and adapt their trajectories to each other to make room for navigation [45]. This model is proposed to overcome shortcomings of models based on anticipate trajectories taking each individual as independent agents that often lead when tested in the wild to ineffective overcautions robot behaviors and even to "freezing the robot" when people attracted by the robot surround it and once the environment surpasses a certain level of complexity, the planner decides that all forward paths are unsafe and freezes in place to avoid collisions. In the case of the "freezing problem", the focus on group collaborative behavior rationales can be more fruitful to design robot's ability to elicit the natural cooperative behavior of making room to create feasibly trajectories. Verbal and nonverbal cues as look at the intended direction or asking for permission could be enough to make room for safe navigation.

Moreover, robots with the "walk around" functionality get involved in spatial relationships with people [2, 12]. Spatial relationships are a combination of distance, relative position and orientation that occur naturally whenever two or more people engage in an interaction [35] and convey significant and relevant social information (e.g. how each of them is involved) and also define an interpersonal space for developing activity.

An interesting approach related to spatial relationships in group walking was conducted by [41]. In this study, the spatial arrangement of 1020 walking groups in an ecological setting was analyzed. In the obtained results several spatial arrangements were observed, depending on the size of the group. Among the observed formations are (see Figure 2.3):

- ‘Line abreast’.- similar to the side-by-side F-formation but walking.
- ‘<-like’ formation (with the walking direction from left to right).- with the middle individual positioned slightly behind in comparison to the lateral individuals.
- ‘Stair’.- an triad arrangement with all members out of alignment, and the central member in an intermediate position between the side members.
- ‘>-like’ formation.- which the central member walks ahead of the left and right members who are aligned to the rear.
- ‘Rhombus formation’.- one person heading the group, followed by a dyad and ended by an another single person.
- ‘1 + 3’.- a single individual followed by a triad.
- ‘3 + 1’.- a triad followed by a single person.
- ‘Arc’.- the central individuals lead the way, followed laterally by the other two individuals
- ‘3 + 2’ and ‘2 + 3’.- in which the group (formed by 5 people) was split in two subgroups.

In other studies [42, 46], Bandini et.al analyzes group behavior such as the characteristics of the groups and their group spatial arrangement while walking in dynamic and crowded environments. In [42] mentions: “At low density, group members tend to walk side by side, forming a line perpendicular to the walking direction (‘line-abreast’ pattern); as the density increases, the linear walking formation turn into a V-like pattern, with the middle individual positioned slightly behind in comparison to the lateral individuals; in situation of high density, the spatial distribution of group members leads to a ‘river-like’ pattern and lane formation, characterized by the presence of a leader that coordinates the group members to cross the space” (see Figure 2.4)

Moreover, [46] analyze different patterns of group spatial arrangement and its significance in relation to the social cohesion of the group. Through experimental and observational methods, the empirical data where analyzed. The results of this analysis are transferred to a synthesis phase, destined to recreate a model of the behavior dynamics of the group.

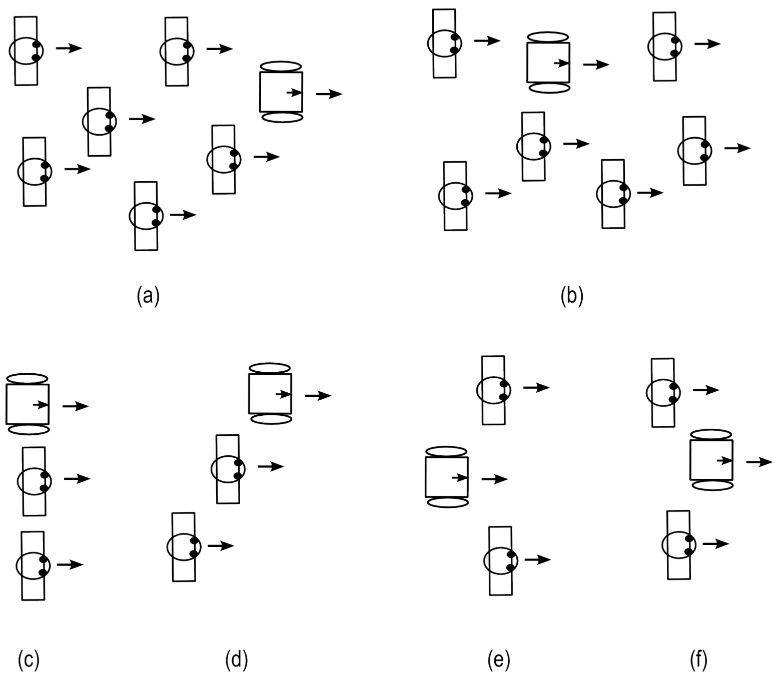


Figure 2.3: Several group walking arrangements: (a) 'Leader-follower'; (b) 'River-like; (c) 'Line-abreast; (d) 'Stair-like'; (e) "<-like"; (f) '>-like'.

In this chapter different concepts and tools have been described that will be useful for the study of human-robot social interactions in real environments. Also different studies on service robot guides in real environments were reviewed. Finally, an approach to evaluate HRI based on spatial relationships was also described.

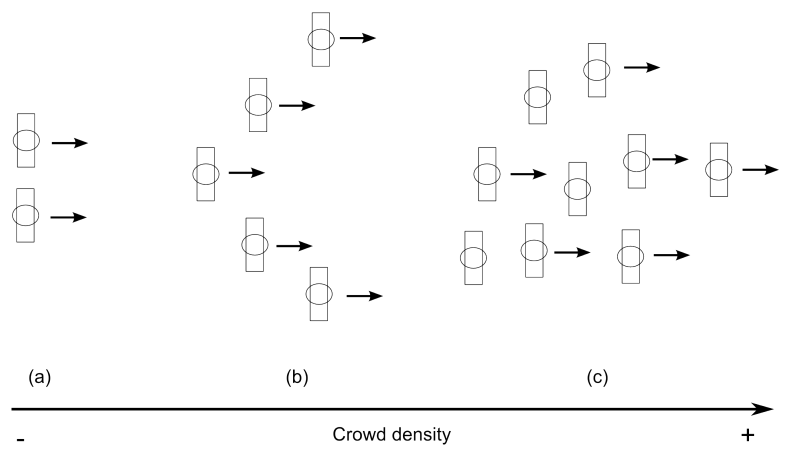


Figure 2.4: Group walking arrangements in crowded scenarios: (a) ‘Side-by-side’; (b) ‘$<$-like’; (c) ‘Leader-follower’.

Chapter 3

Initial Experiences with the Commercial Robot PAL Robotics' REEM at CosmoCaixa

In this chapter we will describe several exploratory studies carried out with a commercial robot in a science museum in Barcelona. These first studies of HRI in an open and crowded environment such as a science museum, will provide clues about the design of our social robot. In these experiences the robot had two types of roles: as a museum guide robot and as an assistant professor robot. Although the role of assistant is not part of our initial model, and also the scenario where it was carried out does not allow us to use F-formations, in this experience an exploratory study was carried out on the social presence and identity of the robot.

In these experiences the robot had two types of roles: as a museum guide robot and as an assistant professor robot. Although the role of assistant is not part of our initial model, and also the scenario where it was carried out does not allow us to use F-formations, in this experience an exploratory study was carried out on the social presence and identity of the robot, two important concepts in human-robot social interaction previously define 2.1.

In Section 3.1 the main features of the robotic platform REEM are described. Next in Section 3.2, a first study about using the REEM robot as a museum guide is introduced. A second experience, with a similar aim, but in a different space of the museum is presented in Section 3.3. Finally, an experience using the robot REEM as an assistant

(a) **(b)**

Figure 3.1: REEM robot from PAL Robotics: (a) REEM's technical data; (b) REEM in a social environment.

robot in an educational session is developed in Section 3.4

3.1 The REEM Robot

From PAL Robotics Company, REEM robot is a 1,65m high semi-humanoid robot with 22 degrees of freedom (see Figure 3.1). The upper part of the robot comprises of a torso with a touchscreen, two motorized arms, which give it a high degree of expression, and a head, which is also motorized. The robot features a rear small platform which can be used to transport objects (e.g. a trolley). The mobile base contains a lithium battery that provides up to eight hours of autonomous operation. A complete range of sensors (i.e. cameras, ultrasonic, lasers) support dynamic distancing and collision avoidance for a safe navigation[1].

As a sophisticated anthropomorphic robot, REEM features several elements and devices to support verbal and non-verbal communication. Some of them are recognizable mechanical versions of natural-like elements as eyes –that are just two holes in the face without lids, eyelids or pupils– framed under the shape of brows. The monochromatic white face presents as well the shape of a nose but no mouth is represented. At both sides of the head are placed two elements evoking vaguely the position and shape of ears that are enlightened when the robot is activated. The head can move up and down and

[1]Visit `http://pal-robotics.com/en/products/reem/` for more details.

Table 3.1: REEM's potential interactive behavior.

Dimensions	Variables	Categories	Subcategories
Verbal	Spoken	Unidirectional	Non conversational
Non Verbal	Gaze behavior	Eye contact	
		Look at	
	Gestures	Head movements	
		Arms motion	
		Hands motion	
	Body stance		
	Displacement	Social navigation	Direction
			Velocity
			Follow
			Guide
			Obstacle avoidance
			Distancing

turn right and left and so does the torso. The articulated arms and hands may support social and utilitarian behavior (i.e. shake hands, point, wave, grasp). In addition users can interact with REEM through the friendly-use 12 inch touchscreen interface on the robot's torso where ad hoc interactive multimedia applications can be run.

Therefore, according to REEM appearance and features enriched intuitive human-like non-verbal communication can be implemented through head and body movement (i.e. gaze behavior), posture (i.e. orientation) and smart navigation (i.e. social distancing) (see Table 3.1). With these types of skills, the robot is able to show different types of behavior, such as speaking, moving the arms, following the faces, guiding, recognizing people, returning to the starting point and speed of approach. This set of behaviors can lead to higher level behaviors, such as: active, passive or sociable.

For safety and feasibility issues the use of some interactive behaviors was deliberately restricted during the navigation of the robot. Consequently the robot's potentiality for verbal and non-verbal communication was reduced to not-facial/not-verbal behavior [6, 47] while the robot is moving. Specifically, arms and hands were blocked and stuck to the body for safety issues. A verbal communication, although not interactive-conversational, is implemented with information displayed on the screen was as well spoken out by the robot as redundant feedback during face-to-face interaction.

3.2 REEM as a Tour Guide Robot. A First Study

This section describes an exploratory study on group-robot interaction with a robot-guide in an open large-scale busy environment. For an entire week a humanoid robot was deployed in the popular CosmoCaixa Science Museum in Barcelona and guided hundreds of people through the museum facilities located in the second floor. The main goal of this experience is to study in the wild the episodes of the robot guiding visitors to a requested destination focusing on the group behavior during displacement. The walking behavior *follow-me* and the face to face communication in a populated environment are analyzed in terms of guide-visitors interaction, grouping patterns and spatial formations. Results from observational data show that the space configurations spontaneously formed by the robot guide and visitors walking together did not always meet the robot communicative and navigational requirements for successful guidance. Therefore additional verbal and nonverbal prompts must be considered to regulate effectively the walking together and *follow-me* behaviors. Finally, we discuss lessons learned and recommendations for robot's spatial behavior in crowded scenarios.

To explore guide robot-visitors performance in open large-scale dense environments, PAL Robotics' REEM robot was deployed during a week in the CosmoCaixa Science Museum informing, motivating, giving directions and walking groups of visitors to requested locations. The whole experience was video-recorded by two external general-view cameras and one on-board camera for observational data analyses. Our approach is to put the focus on the group spatial behavior rather than on individuals taken as independent agents. Therefore, in this study visitors' group behavior while walking (i.e. spatial arrangement) will be described and analyzed –based on the knowledge on group walking and crowd dynamics–, as well as communicative behavior towards the robot.

3.2.1 Study Design

During 6 days the REEM robot was deployed in a restricted area in the CosmoCaixa Museum navigating autonomously around the facility. The robot played the role of a museum guide offering information and guidance to visitors when requested.

The presence of researchers and technical staff was reduced to a discrete and permanent remote surveillance of the robot's performance. The intervention of technical staff was aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

No briefing or instruction was given to visitors and no adaptation of the physical environment was implemented except from the two cameras placed in the walls in an effort to maximize the study ecological validity preserving the natural every-day conditions and routines in the museum activity (i.e. density and flows of visitors).

Objective

The main goal is to observe the robot guiding people (*i.e.* space relationships) and the face-to-face interaction (*i.e.* natural and computer-based interactive behavior).

A first set of research questions are related to robot effectiveness according to its role: Under which conditions and to what extend does the robot attract and entertain visitors, engage people in satisfactory face-to-face interaction, help people to find their way? A second set of research questions are about group behavior during guidance: Can some kind of spatial behavior pattern of the robot group while walking be observed? Are they similar to those modeled for human groups? Are all of them equally suitable for effective guiding? Do visitors' individual variables influence the group walking performance? Do group size and/or composition influence the space arrangements?

The ultimate purpose is to guide the redesign of robot's interactive behavior and –as far as possible in an exploratory study– to draw new knowledge about spatiality and interactive behavior in group-robot interaction.

Scenario and Setup

CosmoCaixa is a science museum located in Barcelona, Spain. The museum hosts a very popular planetarium and a wide range of permanent and temporary exhibitions and attractions where visitors, mainly children and their families, are encouraged to experience and interact actively with the environment. Up to 800,000 people visited the Museum in 2012 [48]. In 2006 CosmoCaixa Barcelona was awarded by the 'European Museum of the Year Award' –institution sponsored by the European Council– as the best science museum in Europe.

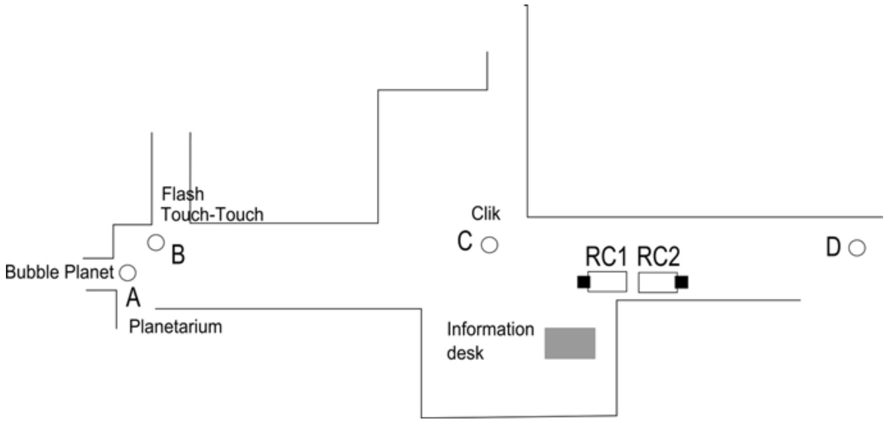


Figure 3.2: Map for the robot placement at CosmoCaixa Science Museum Barcelona

(a) (b)

Figure 3.3: External camera shots: (a) from recording camera 1 (RC1); and, (b) from recording camera 2 (RC2).

The field study was carried out from Tuesday November 27th to Sunday December 2nd, 2012 on the occasion of the European Robotics Week. This 6 days schedule includes a free entrance day (Sunday) when the number of visitors increases considerably.

The robot was deployed in the floor -2 in a restricted area of about 5 meters wide and 40 meters long in a centric corridor leading to the more popular facilities (see Figure 3.2 and Figure 3.3). Three locations (A, B and C in Figure 3.2) were defined as the three possible destinations. Visitors going to "Planetarium" and "Flooded Forest" were walked to A, visitors going to "Flash" and "Touch-Touch!" to B and visitors to the activity "Clik" to C. Point D is the initial location of the robot close to one of the main entrances and besides an information desk.

Task

According to its role the general function of REEM is to enrich visitor's experience by exhibiting itself as an attraction, providing entertainment and information and eventually bringing visitors to requested destinations. The robot role is deployed in three activities: offering services, face-to-face interaction and robot guiding (Figure 3.4).

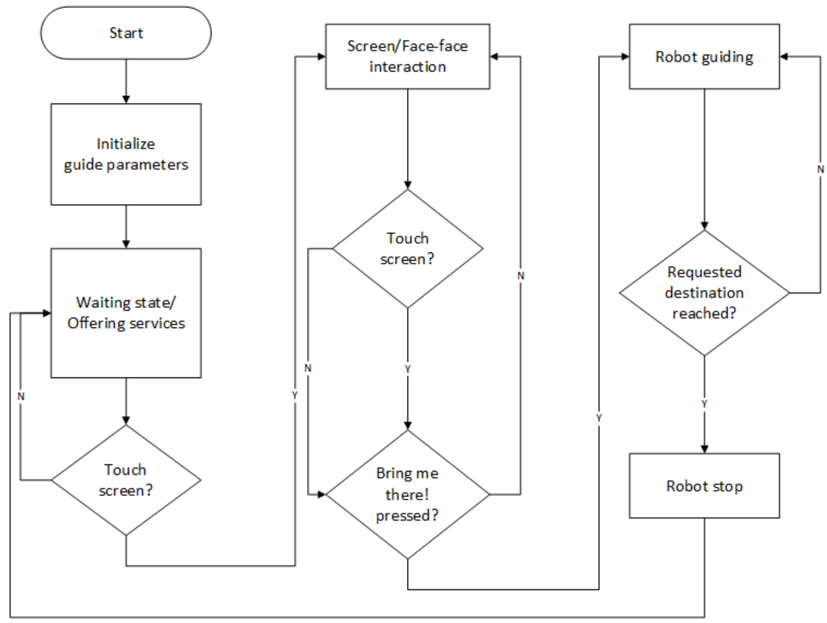


Figure 3.4: Flowchart of robot's guide role.

Different modes of robot behavior were established, as shown in Table 3.2. In passive behavior, for example, the robot does not speak, does not move arms, does not follow faces, does not guide to the requested place, does not know people, does not return to the starting point; while in social behavior the robot speaks, moves the arms, follows the faces, guides to the requested site and knows the people.

Table 3.2: REEM behavior modes

Mode \ Function	Speaks (Y/N)	Arms movement (Y/N)	Face tracking (Y/N)	Guide to the site Y/N)	Meet people (Y/N)	Return to initial point (Y/N)	Approximation speed (+/-)	REEM Alive (Y/N)
Passive	N	N	N	N	N	N	0	N
Active	Y	Y	N	N	N	N	1km/h	N
Sociable	Y	Y	Y	Y	Y	N	3 km/h	Y
Standard	Y	N	Y	Y	N	Y	3 km/h	N

The purpose of offering services is to attract people to engage in interaction. This phase starts as soon as the robot is activated, the ears' lights turn on and the home page is displayed on the touchscreen at the robot's chest (as shown in Figure 3.6a). According to the programmed mode, the robot either deploy a proactive behavior moving around among visitors or remain stationary by the information desk until someone eventually approach. This phase ends when the screen is touched and the interactive multimedia application is launched. During face-to-face interaction the communication is mediated by a graphic user interface, as shown in Figure 3.6. The textual information displayed on the screen is spoken aloud redundantly by the robot to enhance robot's social presence. A tree of the easy-to-use application architecture is shown in Figure 3.5.

If the visitor selects the option *Bring me there!* the robot initiates the guidance navigating to the target location associated to the requested destination. During the guidance there was no interaction between robot and humans. Once the target location is reached or the mission definitive aborted (e.g. the robot is blocked by a crowd) the robot stops and restarts the activity from the first phase (i.e. offering services).

3.2.2 Data Gathering

To register continuously visitors and robot activity two commercial surveillance cameras (RC1 and RC2 in Figure 3.2) were set in the center of the corridor fixed to the building pillars at a height of approximately 3m to have an aerial overview of the experimental area (Figure 3.3). In addition, the robot's on-board camera (RC3) placed behind the robot's eyes was used to obtain a close-up view of visitors from the robot's perspective to study face-to-face interaction. The three video sources were downloaded and stored daily for further processing and analysis.

Observational Data Processing

According to the study's aim the spatial arrangements performed during guidance and visitors' face-to-face communicative behavior with the robot were analyzed. In Table 3.3 the dimensions for group characterization and the coding scheme for group walking behavior and interactive face-to-face behavior are summarized. The coding was carried out manually by two of the experimenters working together.

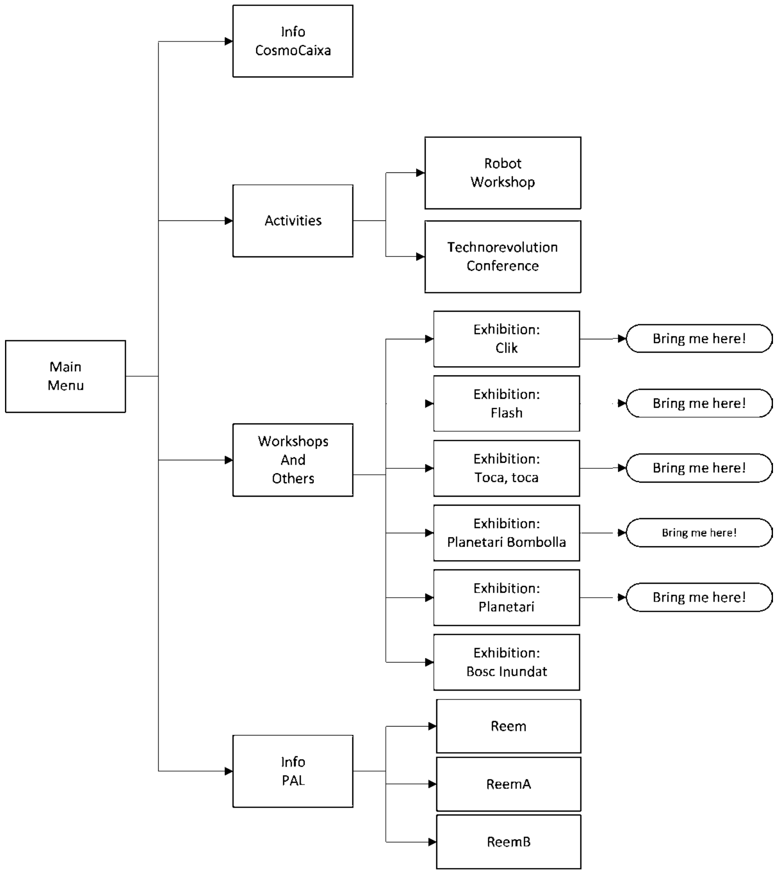


Figure 3.5: Architecture of the multimedia GUI interface.

***Follow-me* episodes** To improve the raw data from the panoramic cameras and to select the relevant episodes of guidance, a preliminary processing of the videos was performed as follows:

- Up to 4828 minutes of recordings from external cameras (RC1 and RC2 in Figure 3.2) were labeled and stored.

- Recordings without any kind of movement were eliminated using computer vision techniques resulting in a total duration of 3966 minutes.

- The sequences where the robot appeared simultaneously with at least one person (i.e. visitors or museum staff) were selected. As a result, a total of 283 scenes with approximately 825 minutes of total duration were selected.

(a) (b) (c) (d) (e) (f)

Figure 3.6: Graphical user interfaces of REEM as a tour guide robot.

- Every episode of guidance, *i.e.* the sequence of group-robot walking together starting when the "Bring me there" option is selected and ending when the robot stops and return to the offer services state, were selected resulting in 91 episodes with approximately 96 minutes of total time.

***Face-to-face* episodes** In total, 83 minutes of video were registered by the camera (RC3 in Figure 3.2) mounted on the robot. From this, 47 minutes of face-to-face human-robot interaction that corresponds to 14 episodes of interaction were selected. For the study of face-to-face behavior an intermittent record of the videos with a sampling

Table 3.3: Group characterization and Visitor's behaviors.

Dimensions	Variables	Categories	Subcategories
Group characterization	Size		
		Single	
		Couple	
		Triple	
		Larger	
	Composition	All-children (Ch)	
		All-young (Y)	
		All-adults (A)	
		Mixed (Mx)	
Visitor's behaviors	Walking Groups	Side-by-side (or line-abreast)	
	Spatial Arrangements	">" formation	
		"<" formation	
		Leader-follower	
	Face-to-Face	Gaze behavior (Gaze B)	Eye-contact (Eye_C)
	Interaction		Look at the robot
			Look at the screen (Look_at Screen)
		Physical contact	Screen
			Other (Oth)
		Facial Expression	Smile
			Grimaces
		Gestures	Wave
			Mimic robot head motion (M_RHM)

frequency of 60 seconds was considered sufficient. Between each sampling an observation window of 10 seconds was considered.

Groups Visitor's groups during the guidance were characterized by age composition and size. Due to the observational nature of the study, age composition were estimated and restricted to the categories shown in Table 3.3. A walking group in guidance is composed by the robot and the visitors that move along with it regardless the relative distance and position between them. Visitors that join the group on the fly –and which probably they don't know where the robot is going– are also considered members of the group.

In face-to-face interaction, social context is defined by all the people that are in camera and within the social space at any moment during the sequence regardless the distance, position, orientation or behavior. Co-present individuals beyond the social distance are only considered when they look at the robot at least once during the sequence.

Group spatiality during guidance In this study a guidance episode is the sequence of walking together behavior deployed by the robot and a group of visitors starting when the “Bring me there” option is selected on the screen and ending when the robot stops and comes back to offering services state. Guidance ends either when the robot reach the requested destination or when the robot’s trajectory is definitely aborted (e.g. robot blocked by a crowd of visitors, robot stuck in a corner, emergency shutdown). Temporary stops during the displacement due to navigational constraints (i.e. mobile obstacles avoidance) do not end the guidance sequence provided the trajectory is resumed by the robot. Any displacement of visitors along with the robot is considered a guidance episode regardless to the particular relative position they adopt (i.e. robot ahead, robot side-by-side).

The size categories in guidance sequences are single individuals, couples, triples and larger groups (see Table 3.3). Group composition is referred to the age of the group members defining four types of group: all-children, all-young, all-adults and mixed groups.

An ad hoc coding scheme was built-up to investigate the space distribution patterns –relative position and distance between agents including the robot. Four spatial patterns were described to classify the group spatial arrangement while walking with the robot: side-by-side or line-abreast, leader-follower, “<”-like and “>”-like (see Table 3.3 and Figure 3.7).

Visitors-Robot Face-to-face Interaction

According to the task description, face-to-face interaction may happen any time the robot is activated and not engaged in guiding a group. In this situation when typically the robot is stationary the touch screen-based interaction is available and REEM’s social behavior is based on head motion and speech.

The 47 face-to-face sequences are described according to participant individual variables (genre, age), presence of other co-located visitors (group size and composition), the distance from the robot, the robot behavior (displacement, head motion) and visitor social behavior: gaze (eye contact, look at the robot, look at the screen), physical robot

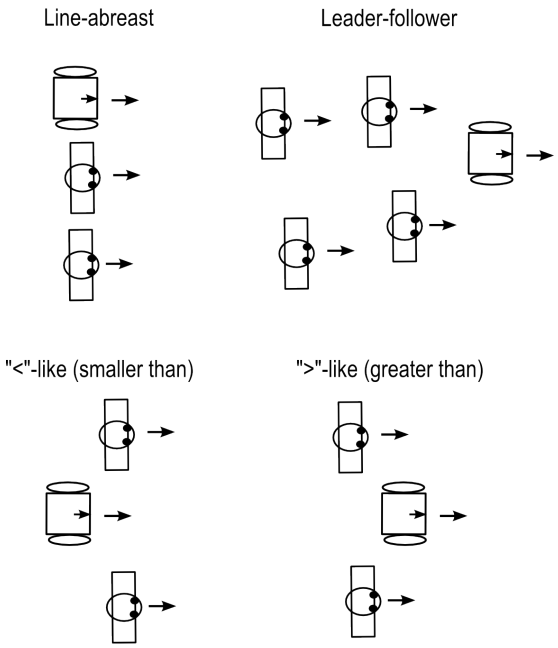


Figure 3.7: Spatial arrangements described in this study.

contact (on the screen or other), facial expressions (smile, grimaces) and gestures (wave, mimic robot head motion) (see Figure 3.8 and Table 3.4).

3.2.3 Results

During the 6 day trial the robot completed 48 hours of autonomous operation walking through the museum defined space during regular public attendance at a maximum displacement speed of $4km/h$. During this time REEM robot operated without remarkable pauses (i.e. more than one hour) and suffered a total of 5 shutdown incidences caused by visitors' misuse pressing deliberately the emergency stop button placed at the robot's back.

Group-robot Description

From the whole *follow-me* episodes analyzed, 11.64% where people and robot together in the scene. Observing the group-robot composition, 1.10% of groups that interact with the robot were all-children, 8.79% were all-youth, 52.75% were all-adults and 37.36% were mixed groups, as shown in Figure 3.9a. From the mixed groups, 50% were formed

Table 3.4: Face-to-face interactive sequences description.

									Gaze B			Contact		Facial		Gestures	
Sample	Subject	EPISODE	Age	Genre	Size	Comp	HM	DIS	Eye_C	Look_at	Screen	Screen	Oth	Smile	Grimace	Wave	M_RHM
S02	V_01	B	Young	F	10	y	Y		1					3	0	0	0
S03	V_02	B	CH	M	10	y	Y		3		2			3	1	0	0
S04	V_03	B	Young	M	10	Y	Y		1					0		0	0
S05	V_04	C	A	M	1	A	Y		1					0		0	0
S06	V_05	D	A	F	2	A	N		0					1		0	0
S07	V_06	D	A	M	3	A	Y		0		1	1		1		0	0
S08	V_06	D	A	M	3	A	Y		1		1	1		0		0	0
S09	V_07	D	A	F	3	A	Y		2		1	1		0		0	0
S10	V_08	E	A	M	2	A	N		1					0		0	0
S12	V_04	F	A	M	1	A	N		0		1			0		0	0
S13	V_09	G	CH	M	30	Ch/Mx	Y		5					1		1	0
S14	V_10	G	CH	F	30	Ch/Mx	Y		3		1			2		3	0
S15	V_11	G	CH	M	30	Ch/Mx	Y		3		1			1		1	0
S16	V_12	G	CH	F	30	Ch/Mx	Y		5					1		1	0
S17	V_13	G	CH	F	15	Ch/Y	Y		3					2		1	0
S18	V_14	H	A	M	1	A	Y		1					1		1	0
S19									0					0		0	0
S20	V_15	H	A	M	1	A	Y		1					1		0	0
S21	V_16	I	Young	M	5	Y	Y	y	0		1	1		0		0	0
S22	V_17	J	A	F	1	A	N		1					1		0	0
S23	V-18	J	CH	M	6	CH/A	Y		1					1		0	0
S24	V_19	J	Young	M	3	Y/A	Y		0		1	1		0		0	0
S25	V_20	J	A	F	3	A	N		1					0		0	0
S26	V_17	J	A	F	6	mx	y		1					0		0	0
S27	V_21	K	A	F	1	A	y	Y	1					0		0	0
S28	V_22	J	A	M	1	A	Y		1					1		0	1
S29	V_23	J	Young	M	4	Y	Y		1					1		0	1
S30	V_24	J	A	M	7	MX	Y		1		1	1		1		0	0
S31	V_25	J	CH	M	6	MX	Y		2		1			0		0	1
S32	V_26	J	Young	M	7	MX	Y		1					1		0	1
S33	V_27	J	Young	M	7	MX	Y		1		1			1		0	1
S34	V_28	J	Young	M	11	MX	Y		0		1	1		0		0	0
S35	V_29	L	CH	M	5	MX	Y	Y	0	1				0		0	0
S36	V_30	J	CH	f	8	mx	y		2					2		0	1
S37	V_31	J	CH	F	13	MX	Y		0	9				1		0	0
S38	V_32	J	A	M	6	MX	Y		0	1	5			1		0	0
S39	V_33	M	CH	M	6	MX	N		0	10				0		0	0
S40	V_34	M	A	F	6	mx	Y		0	8				2		0	0
S41	V_35	J	A	F	10	MX	Y		0	7				0		0	0
S42	V_36	J	A	F	8	MX	Y		0	8				0		0	0
S43	V_37	J	A	F	11	MX	Y		0	7				0		0	0
S44	V_38	J	A	m	3	Mx	Y		1		1	1		0		0	0
S45	V_21	J	A	F	6	Mx	Y		0		1	1		0		0	0
S46	V_39	N	A	M	1_2_4	A_Mx_mx	Y		2					0		0	0
S47	V_40	N	CH	F	5	Mx	Y		0		1			0		0	0
S48	V_41	J	CH	M	2_3_4	ch/ch/Mx	Y		0		1	1		0		0	0
S49	V_42	O	A	F	1	A	Y		2					1		0	0

Figure 3.8: Face-to-face HRI examples. Facial expressions: (a) smile and (b) grimace; and gestures: (c) wave and (d) mimic robot head movement.

by children and adults, 29.41% were formed by youth and adults, and 20.59% were formed by children, youth and adults, as shown in Figure 3.9b.

Concerning the size, 3.30% of the people walked alone with the robot, while the 96.70% walked in groups: 10.99% of groups were couples, 14.29% triples and 71.43% larger groups, as shown in Figure 3.10a and 3.10b.

Relating group size with its composition, we see that the visitors who walked alone with the robot were 33.33% youth and 66.67% adults; couples were 10% youth, 80% adults and 10% children and adults; triples were formed by 15.38% youth, 61.54% adults, 7.69% children and adults, and 15.38% youth and adults. Larger groups were composed of 1.54% children, 6.15% youth, 46.15% adults, 23.08% children and adults, 12.31% youth and adults, and 10.77% children, youth and adults, as shown in Figure 3.11.

Group-robot Spatial Arrangement

In the present study we focus specifically on the *follow me* behavior. In the context of guidance, a *follow me* episode is defined as a walking group formation including the robot displacing together and forming typically any of these two spatial arrangements: *walking along with* (*i.e.* robot and people aligned side-by-side perpendicularly to the

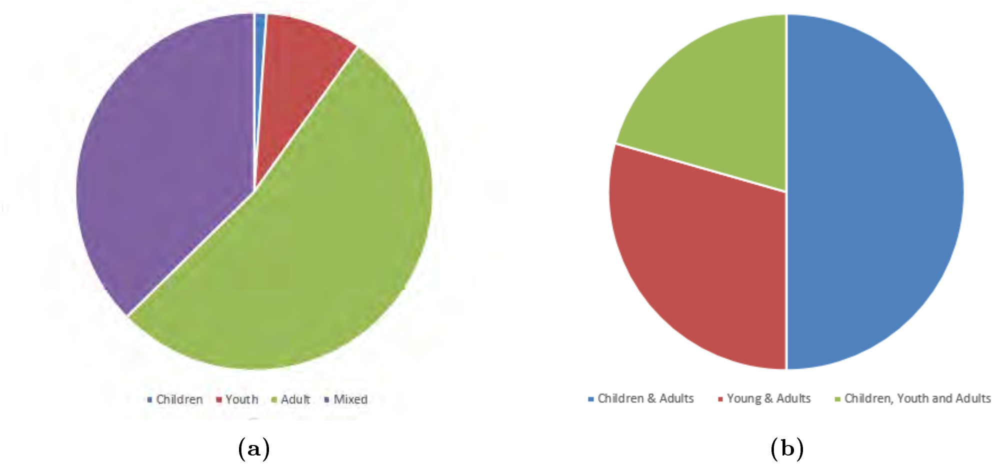


Figure 3.9: Group-robot description: (a) Group composition; (b) Mixed group composition.

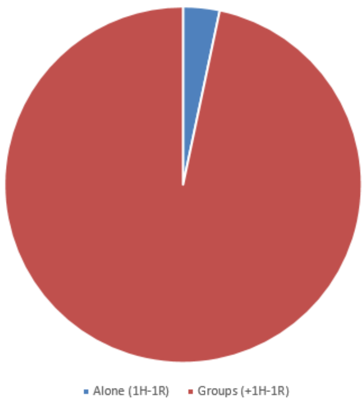
Table 3.5: Spatial relationships by group's size.

	Group size									
Formation	1	2	3	4	5	6	7	8	9+	Total
Leader-follower	3	8	13	4	9	2	6	4	33	82
Side by side	0	2	0	0	0	0	0	0	0	2
Unidentified	0	0	0	1	1	0	1	0	4	7
Total	3	10	13	5	10	2	7	4	37	91

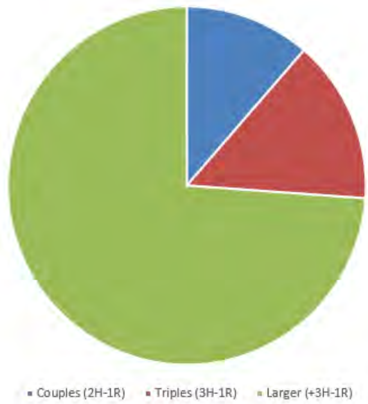
walking direction)(as Figure 3.12a) or *walking ahead* (i.e. visitors follow the robot) (as Figure 3.12b).

91 episodes of *follow me* were observed. In almost all the cases a *walking ahead* formation was observed (see Table 3.5), where the robot leads the group and visitors follow behind. While this space arrangement is a natural disposition for following a leader for high density groups [42], this type of formation pose important constraints to communicate and therefore to generate social cohesion between the robot and the visitors during guidance. On the other hand, only 2 side-by-side spatial formations were observed during guidance and in both cases it was just two people walking with the robot.

Results about group spatial arrangement with people walking together with robot in the *follow-me* behavior showed that:



(a)



(b)

Figure 3.10: Group-robot description: (a) Alone Vs Group (b) Group size.

- 100% of guide-visitor couples (i.e. one person-one robot) was characterized by a leader (robot) - follower spatial arrangement, as shown in Figure 3.13a;
- 90% of guide-visitors triples was characterized by the robot heading the group and followed by a dyad in an ">"-like pattern (see Figure 3.13b), and 10% by "<"-like pattern (see Figure 3.13c);
- 100% of four-agents groups (*i.e.* 3H-1R) was characterized by the robot followed by a triad (see Figure 3.13d).

I's observed that 96.15% of the formations analyzed have a robot leader- human follower formation, indicating a weak social cohesion between the robot and people in almost all spatial arrangements. Two can be the causes of this type of training, one that the robot does not perform any type of interaction during the guidance, so there is no motivation on the part of the human to undertake a dialogue; the other possibility has to do with the physical scenario where the guide is developed, the same that takes place in a narrow corridor in comparison with other sites of the museum.

Face-to-face HRI

As mentioned before, a subset of 47 video samples (10 seconds of sampling in each minute) of face-to-face interaction from the on-board camera were analyzed, whose results can be seen in the Table 3.4. From these samples, 42 subjects were observed

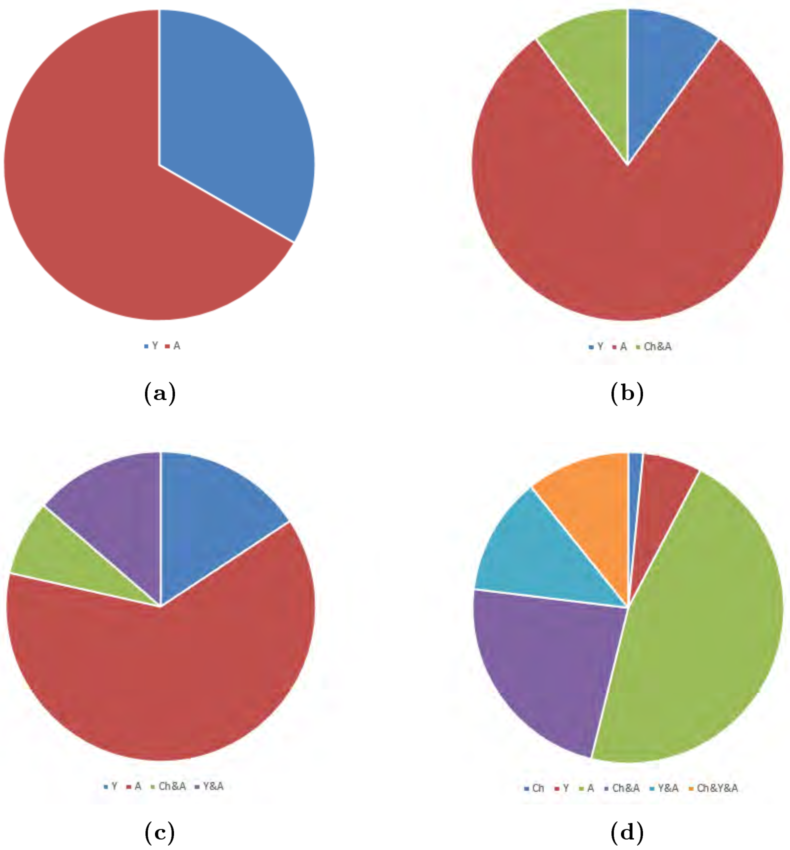
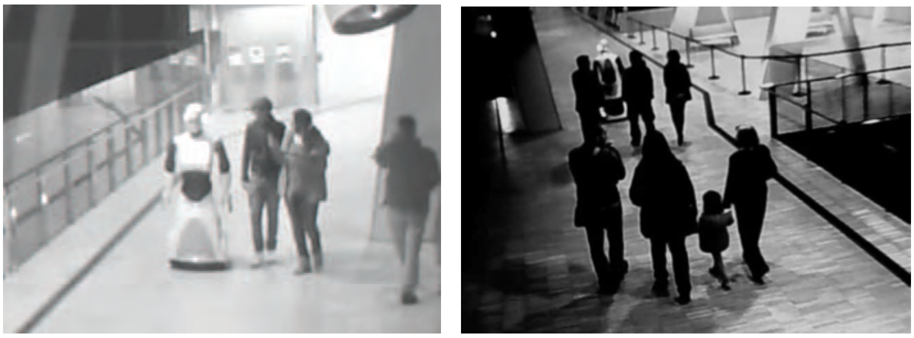


Figure 3.11: Group-robot size and ages composition: (a) Alone; (b) Couples; (c) Triples; (d) Larger.

interacting with the robot as well as 14 episodes or scenes of the same set of people in the interaction. The table also shows the age and gender of the interacting partner, as well as the size and composition of the group. In a separate column it is indicated if the head of the robot performs some kind of movement (HM DIS). The meaning of the acronyms used in this table can be found in the Table 3.3. You can observe the results of the different types of behavior of people, such as gaze behavior, physical contact, facial expressions and gestures.

Thus, among the 47 sequences, 60% of the 43 participants were males and 30% females. According to the age, 45% were adults, 17% young people and 30% children.

In 29 out of the 47 sequences (62%) the visitor established eye contact with the robot at least once (Figure 3.14a), with a total of 50 smiles registered. In 9 occurrences



(a) (b)

Figure 3.12: *follow me* spatial arrangement types: (a) *walking along with* and (b) *walking ahead*.

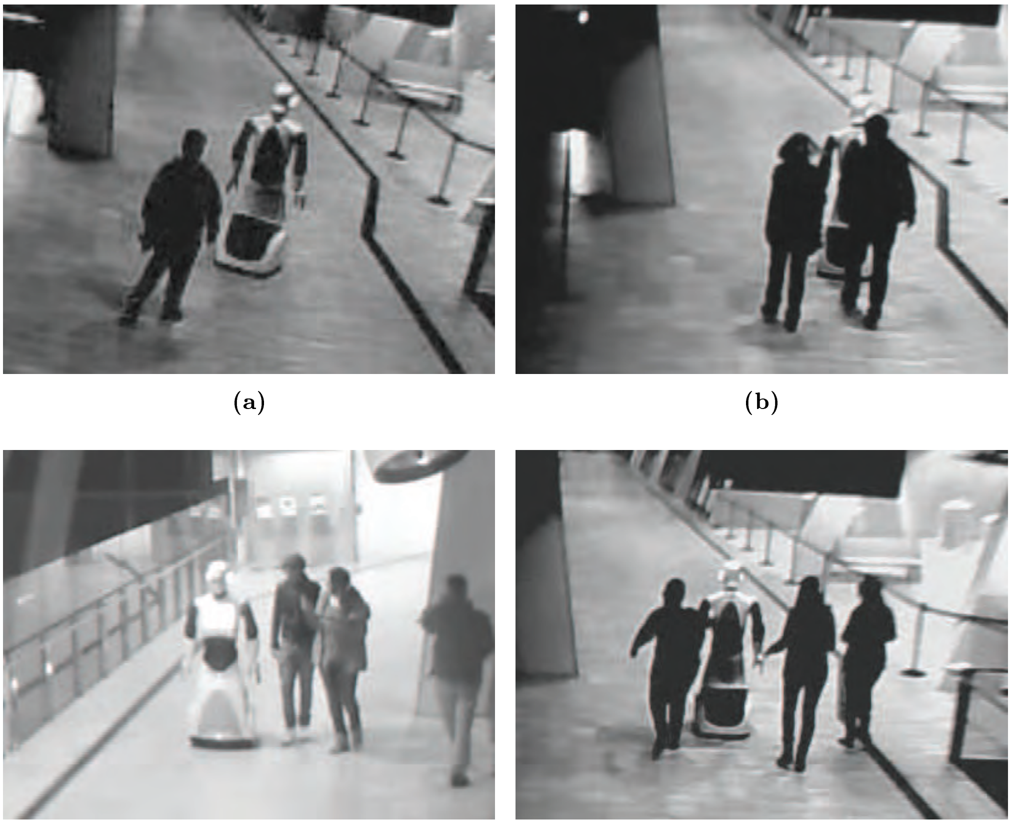
the visitor kept staring at the robot's face during the whole sequence.

In 23 sequences (62%) the visitor smiles at least once with a total of 31 smiles registered. In 9 occurrences the visitor kept smiling during the whole sequence (Figure 3.14b).

With respect to the interface mediated communication, in 10 sequences (21%) the visitor touched the screen at least once and in 18 (38%) the user looked at the screen. Male visitors seem more prone to interact through the graphic screen: 50% of male visitors looked at the screen and 29% touched the screen, in comparison with a 22% of female visitors that looked at the interface and only 11% of females touched it.

From 8 sequences showing people interacting alone with the robot (Figures 3.14(a)–(b)), in 7 cases the visitor established eye contact and in 5 sequences they smile (see Figures 3.14(e)–(f)).

It is noteworthy that the analyzed behaviors are not mutually exclusive. Even though in the descriptive analyses are quantified as independent behaviors, facial expressions (e.g. smile) and gestures (e.g. wave) may be presented simultaneously and actually usually are in non-verbal interpersonal communication (e.g. 100% of observed wave behavior appeared with smile).



(a) (b) (c) (d)

Figure 3.13: Group spatial arrangements: (a) guide-visitor couple with leader-follower formation; (b) guide-visitors triple in a leader-follower spatial arrangement with robot heading the group followed by a dyad; (c) guide-visitors triple with “>”-like pattern; and, (d) four-agents with robot leader followed by a triad.

3.2.4 Discussion

Robot Guide’s Performance

Initially in this study a schedule with different interaction modes of the REEM robot had been established such as: passive, active, social and standard; where each modality was defined by a set of skills such as speaking, movement of arms, face tracking, guide to the site, meet people, return to initial point and approximation speed. However, this was not possible, either due to the company’s need to carry out the best experiences of interaction with the user, as well as the complexity to modify these modes in-situ.

It was observed that on several occasions the robot performed the role of a museum

(a) (b) (c) (d) (e) (f)

Figure 3.14: Interactive behaviors: (a) visual contact; (b) smiling; (c)–(d) waving; (e)–(f) mimic head movement (faces have been blurred)

guide robot -attracting people, providing information and guidance- fulfilling up to 91 *follow-me* sessions operating autonomously. Registered face to face episodes show visitors spontaneous social behavior addressed to the robot including eye contact, smiles, and greetings. Although visitors' face-to-face behavior is described qualitatively and no conclusive results can be drawn from the data some of the behaviors observed show the enjoyment and engagement of the people.

However same shortcomings and difficulties were met mostly related to the challenging social context that is extremely complex, dynamic (i.e. changeable in visitors density and distribution) and sometimes crowds of visitors with uncertain trajectories. Attendance is formed by a wide range of visitors' profiles and group configurations with a high rate of children that increases the uncertainty. The physical scenario is not stable either due to temporary exhibitions and events and maintenance tasks [16].

In addition, a science museum as an experimental bed test for HRI studies has some peculiarities that may be outlined. The robot is an attraction itself as a piece of smart technology and an object of visitors' interest and curiosity in a context where visitors are encouraged to explore and try [11]. Far to become transparent in this situation the technology becomes the target and visitors do not miss the opportunity to explore and interact with the robot manipulating it (*e.g.* pushing the emergency button), defying its capabilities (*e.g.* climbing to the rear platform) and putting it in challenging situations to see what happens (*e.g.* activating on purpose the face tracker moving the head up and down). In our trial we have observed that eventually these visitors' active behaviors which results in a system shutdown or the impossibility to fulfill the task.

The robot guide attracted untrained naïve people and engaged visitors 91 times in *follow-me* behavior without any other cue but the robot appearance and behavior, especially when it moved around, moved the head searching and tracking faces and initiated motion. The success in attracting people led the guide robot to face the "freezing robot" problem: once the environment overpasses a certain level of complexity all alternatives are unsafe and stuck in place [45].

Therefore, the context of service is a challenging combination of a complex space and the willingness of people -sometimes crowds- to approach and interact with the robot. In this situation a conservative navigation for safety issues and a focus on robot's robustness is required even if it implies a sacrifice on robot's interactivity and attractiveness (e.g. discarding the communicative use of its arms). From this experience, we agree with Willeke [17] that resilience for recovering from visitors' misuse -even abuse- and awkward situations (i.e. approaching a wall too closely or being crowded by people) is a crucial issue to ensure the continuity of the service.

To overcome these constrains, we consider that it would be interesting to empower the robot with some kind of authority that would be consistent with the role of guide to regulate visitors' behavior (i.e. showing people clearly what is not allowed) and to give the robot more social presence that maybe could prevent from same rough manipulations. In addition, closer but unobtrusive supervision might be provided by the staff to discourage deliberate or not deliberate misuse.

Contributions to HRI Studies

The open environment due to the few changes in the regular museum schedule and in the physical scenario - except for the 2 cameras mounted in the walls - allowed to perform a non-participant observation, that is, the subjects were not aware of the presence of the observer. Thus, the activity of subjects not affected by bias, *i.e.* had a null reactivity. In spite of this, it is necessary to take into account the spaces where the interaction will take place, especially the location of the cameras that should cover the entire study area. However, this represents a challenge on large surfaces, as in this study.

The literature on social space arrangement, crowd behavior and spatiality in walking groups are revised and applied to the evaluation of hybrid groups formed by the robot and a number of naïve participants. This approach focuses on groups rather than on individuals and extend the scope of HRI proxemics -mostly oriented to one-to-one interaction [2]- with the consideration of dynamic spatial arrangements during displacement that are critical for robot's performance in public spaces.

The detailed description of the HRI episodes offers empirical based insights that can be of interest to improve the evaluation and design of HRI in public and dense environments and to focus on relevant variables like the social situation, the role the robot's takes, its social affordances, the actual robot's behavior and the physical constrains.

Furthermore, this study has outlined the feasibility and convenience of automatic processing techniques (e.g. computer vision) to study spatial HRI through systematic observation even though their use in-the-wild presents great challenges of reliability and robustness .

Limitations and Future Work

Several lessons learned to be taken into account for future exploratory studies include:

Filtering the video sequences There is an effort of filtering video sequences of the robot guiding people. As a solution we can extract the videos from the moment when the option "bring me there" is selected. To do this, you should have a log with the times of these events, which then allow you to search with more precision and less workload the guidance videos.

Reliability Although the observational study was carried out by two observers, an analysis of the reliability was not performed and therefore the observations can not be validated. In spite of that, it has been believed that it has been considered interesting to add it for completeness of the cross-sectional studies carried out.

Coverage The two cameras do not cover in detail all the relevant interactive behaviors between robot and visitors in the experimental space. In particular, even though the perspective provided was enough for group arrangements classification, face-to-face interaction at the end of the guidance episodes are missed because the destination points were too far away from the cameras. Adding another on-board omnidirectional camera could provide a view of all the social space around the robot and facilitate the use of automatic spatial behavior analyses.

Face-to-face interaction data Due to technical constraints it was not possible to videotape continuously face-to-face interactions from the on-board subjective-perspective camera so the observation has been done on a sample of sequences from the 83 minutes of available recordings. No systematic or representative conclusions can be drawn from the results. However, the analysis of the observed data gives us some clues to improve systematic observation in future research. As for example observing and comparing the behaviors of both the robot and the visitor, one could relate the emotional state of the robot with the visitor's gaze, facial expressions and gestures.

Measures of density A very relevant variable that influences group walking behavior is the space density, which has not been measured accurately in our study. Further studies must provide density measures preferably automatically obtained from video processing.

Logging of robot behavior and interaction on GUI To triangulate human-robot interaction evaluation it would be very interesting to analyze the logs of robot behavior and of the interface-based interaction. Data are in principle available from robot's log but it must be faced the issues of processing and storage.

Visitors' experience The evaluation of the guide robot performance would benefit from a complementary assessment of visitors experience from short questionnaires after interacting with the robot. This self-report data could be of great interest to interpret or contrast the observational data. This however presents logistical challenges for these natural environments.

Factors influencing behavior While the aim of this study is to provide descriptive analyses and evidence-based insight with high ecological validity, systematic exploration of factors influencing the observed behavior should be carried out in future works to better understand the interactive behavior in this context. Systematic quantitative studies on human variables (like gender or age), social context, robot personality (like body language), and/or group behavior (like spatial arrangement, composition, density, dispersion or velocity) could be of interest in social HRI research.

3.2.5 Conclusions

An exploratory study on group-robot interaction was carried out during a week in an open and natural environment to observe visitors' spatial behavior and communication with the guide robot REEM in a popular science museum.

The robot succeeded in developing the role of a museum guide -attracting people, providing information and guidance- fulfilling up to 91 *follow-me* sessions operating autonomously. Registered face-to-face episodes show untrained visitors social behavior addressed to the robot including eye contact, smiles, and greetings.

Differently from previous works on mobile service robots that evaluate navigation and HRI as separate functions we address the study of the spatial behavior while navigating, focusing on its social meaning, not only as a prerequisite for effective communication (i.e. orientation, positioning) but as potential communicative acts (i.e. express intent and emotions).

The analysis is focused on visitors' groups rather than individual. Groups were described according to their composition, size, spatial formations and interactive behavior with the robot during guidance. Observational methods applied to evaluate group-robot interaction provide fruitful insight to understand the relationship between robot positioning and efficient communication (*i.e.* . walking side-by-side) and between robot motion cues (*e.g.* gaze behavior, body orientation) and collaborative walking together behavior through populated environments.

From literature and from the observations gathered during the field test we can indicate that constraints in spatial arrangements can affect the social field for shared activity and thus can influence the quality of human-robot interaction. Restrictions regarding the physical scenario and when the lack of interaction of the robot when walking, can lead to contradictory spatial formations (*i.e.* a small group of 2 people having a leader-follower spatial formation).

appropriateness of the spatial relationship is a key issue in collaborative activities as guidance that requires continuous communication (i.e. offering to bring, heading a destination) and mutual regulation (i.e. group walking together).

Lessons learned from this long lasting study could also be considered for designing spatial behavior of service robots in other contexts as hotel receptions, leisure parks or hospitals.

Spatial arrangements during guidance may not be effective when confronted with robot's affordances and navigation constraints and therefore limit the system performance. However, social robot behavior can be improved by the robot through new forms of verbal or nonverbal communication.

3.3 REEM as a Tour Guide Robot. A Second Study

In this section we describe a second experience using the REEM robot as a museum guide. This activity was developed on the last weekend of November 2013 in the floor

−5 of the CosmoCaixa Science Museum in Barcelona. In comparison with the first observational study, some differences exists about the scenario, role of the robot, and data analysis, which together with the lessons learned from the first experience, should be taken into account as far as possible.

The present study focuses on the description of both face-to-face and walking behaviors. The face-to-face behavior refers to the social interaction that occurs when the group-robot is static, that is, it does not move; while the walking behavior when the group-robot is walking or moving. Data were gathered in a naturalistic way and analysed applying using proxemic, group walking [41] and crowd dynamics [46] theories.

From here, the study design is described in Section 3.3.1. Next, the data analysis were explained in Section 3.3.2. Then the the obtained results are presented in Section 3.3.3. The discussion and conclusions from the study are finally developed in Sections 3.3.4 and 3.3.5 respectively.

3.3.1 Study Design

For two complete days the REEM robot was deployed in the CosmoCaixa Science Museum navigating autonomously around the museum. The robot played the role of a museum guide offering information and guidance to visitors when requested.

The presence of researchers and technical staff was reduced to a discrete and permanent remote surveillance of the robot's performance. The intervention of technical staff was aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

No briefing or instruction was given to visitors and no adaptation of the physical environment was implemented, thus preserving the original conditions and services provided in the museum.

Objective

The main objective in this Section is to study human-robot interaction with REEM robot as a museum guide robot.

More specifically, we want to characterize the social interaction of both the group-robot that do not move (*i.e.* face-to-face behavior) and the group-robot that are in movement (*i.e.* walking behavior) along the guided route.

Although in this study REEM is a museum guide robot, there are certain differences with the previous study.

Scenario and Setup

The field study was carried out also in the CosmoCaixa Science Museum of Barcelona during the weekend at the end of November 2013. This time the robot was deployed in the floor -5 in a wide area surrounded by several permanent exhibitions. The tour carried out by the robot consisted of 6 stops (see Figure 3.15), being 'Stop 1' (near the Foucault Pendulum) the starting point and 'Stop 6' (near the Flooded Forest) the arrival point of the tour, in each of which the robot stopped and gave information about the museum (see Figure 3.16).

Task

Unlike the previous study, in this study the museum guide robot has a predetermined route, which does not depend on the selection of the user. From the starting point (Point 1: Foucault Pendulum) to the point of arrival (Point 6: Flooded Forest), the robot goes always through a series of stops where it explains part of the tour, as seen on Figure 3.16.

In the previous study, the guidance of a group of visitors was only from one initial site to another arrival. Also the different points of arrival were very close and there was almost no distinction from each other. In this case it is expected that the same group of visitors accompany the robot at the different stops.

3.3.2 Data Gathering

Given the large area to be covered in the study, placing a set of cameras recording the whole guidance was technically unfeasible. Instead, a single observer took photos from different points of view along the route. In order to minimize the reactivity, as well as to have an aerial overview of the experimental area, the observer was located on the floor -4 of the museum, as shown in Figure 3.17.

From here, the observer arranged a series of points in which he would make the shots in order to obtain the greatest amount of visual information during the interaction (as

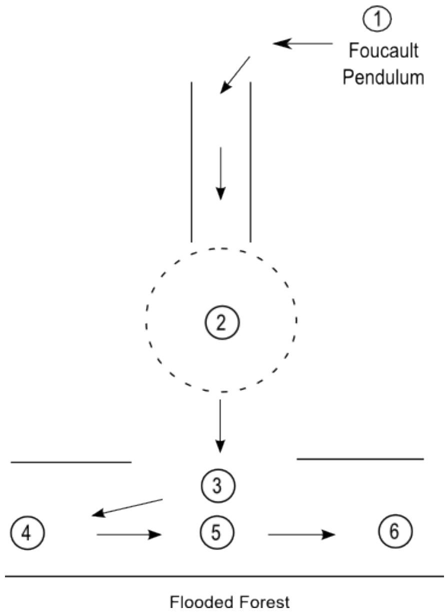
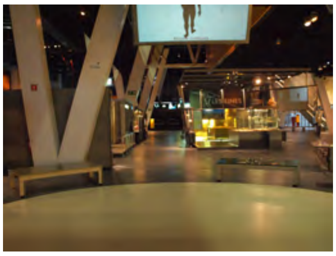


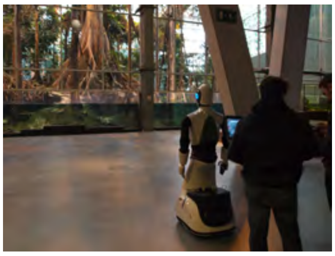
Figure 3.15: Tour guide robot map setting

(a)

(b)



(c)



(d)

Figure 3.16: Several camera shots during the robot setup: (a) REEM robot at the first Stop; (b) Robot walking in a corridor; (c) REEM robot at the Stop 2; (d) Robot in Stop 3 (see the Flooded Forest at the background).

shown in the Figure 3.18). To record the times, the time at which the photos were taken was used; even though those times only had a resolution of minutes.

The entire observation area was segmented into different locations, based on the objectives set. In this way two types of spaces were defined: those spaces where the robot remains still (points 1, 2, 3, 4, 5 and 6) which are the start (point 1) and the arrival (point 6) points, as well as the intermediate information points (points 2, 3, 4 and 5); and those spaces where the robot is navigating, that is, the transit spaces between two contiguous points (1-2, 2-3, 3-4, 4-5 and 5-6). The transit from point 6 to point 1, *i.e.* the return of the robot to its starting point, is outside the scope of the present study (see Figure 3.16).

In this study the group behavior around the robot was studied both by their size and by their spatial arrangement. Two category systems were built-up both to register both of them (see Table 3.6). As can be seen in the table, we consider the spatial arrangements that may exist in both face-to-face interactions when the human-robot group is fixed and in interactions when the human-robot group is walking. In addition, with respect to the previous study, the types of spatial arrangements (such as the "stair" arrangement) have been extended for a more detailed analysis of the formations.

Group Size

During the first study the size of the group was referred only to people around the robot. Excluding the robot from the group generated some debate on whether or not the robot should be considered part of the group. Therefore, for this study it was defined the *group-robot* as the set of humans and robots that are together showing a social interaction. In this study, a group-robot may be composed of a dyad, when the group is formed by one person and one robot (1H-1R); a triad when the set consists of two humans and one robot (2H-1R); a 4 group-robot will correspond to a set of 3 humans and a robot (3H-1R) and an large group-robot corresponds to the set formed by people > 3 and a robot ($>$ 3H-1R). All these representations are depicted in Figure 3.19.

Spatial Arrangements

The spatial arrangement category describes the group spatial arrangements during the guiding service. These categories are separated into two large groups: the categories that correspond to the F-formations (see Figure 2.2), that is, those face-to-face formations

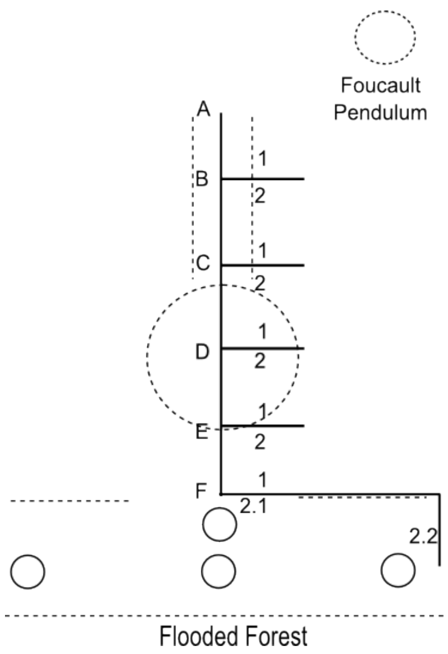
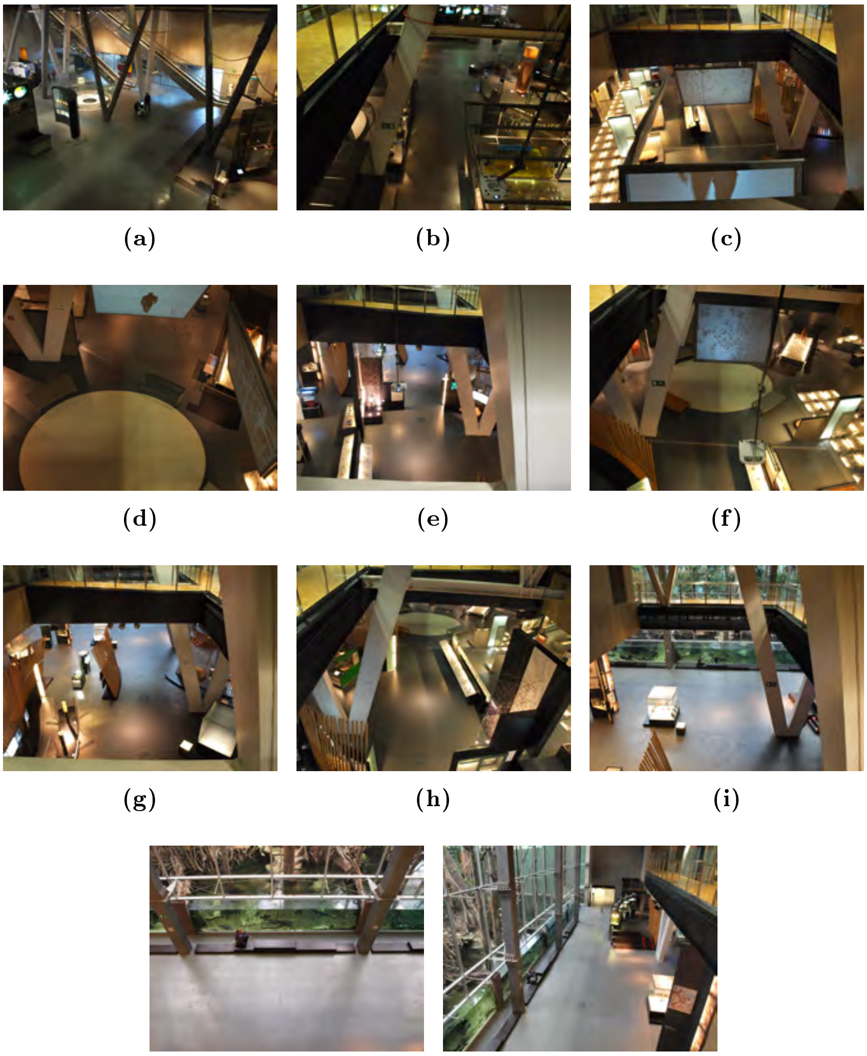


Figure 3.17: Observer map setting

Table 3.6: Group size and spatial arrangement categories

Dimension	Variables	Categories
Group size		Dyad
		Triad
		4-Group-robot
		L-Group-robot
Spatial Arrangement	F-Formations	Vis-a-vis
		L-shape
		Circular form
		Side-by-side
		Horseshoe shape
		Cluster
		Performer-audience
	Walking Groups	"<"-like
		">"-like
		Stair
		Line-abreast
		Leader-follower
		River-like



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure 3.18: Observer setting camera shots during the setup: (a) View from point A1; (b) View from point B1; (c) View from point B2; (d) View from point C1; (e) View from point C2; (f) View from point D1; (g) View from point D2; (h) View from point E1; (i) View from point E2; (j) View from point F1; (k) View from point F2.

in situations where all individuals are not moving; and the walking formations (see Figure 2.3), which occurs when the whole group is walking. It should be noted that there are formations where one of the subjects has a preponderant role, as in the case of a performer-audience or the leader-follower. In this study, it is expected that the REEM robot fulfills this role.

The group behaviors were recorded in check sheets for each session, taking into account that one session includes all the guidance interaction from point 1 to point 6.

3.3.3 Results

During the two days of the experiment, a total of 11 sessions which represents 221 minutes were observed. An average time of 20:05 minutes per session and a standard deviation of 6:26 minutes. Likewise, a total of 113 spatial arrangements and group sizes were observed. The detail about the location times, group sizes and spatial arrangements is given below:

Location times

According to the results shown in Table 3.7, site 1 corresponds to the location where the robot was the highest amount of time during the guidance: 4:11 minutes on average and a standard deviation of 4:25 min. It should be mentioned that this location, as well as the site 6, have the highest standard deviations, which is consistent with the variability times that are usually given at the beginning and at the end of the guidance, since both times during stops 1 and 6 depends on when a visitor initiates the guidance through the touchscreen or when the operator gives the order to the robot to return to the base, respectively. See Figure 3.20, in order to help visualization of data and easier to observe trends of timings at each spot.

Group size

From the 113 group-robot sizes observed, 81% corresponds for Large group-robot, while a 9% goes to 4 Group-robot, 5% to triads and 4% to dyads (see Figure 4.14).

Moreover, Table 3.8 summarizes the relative frequency of group size observed at different locations. The relative frequency corresponds to the percentage of each group of the total groups observed at a certain point. The number of people computed at

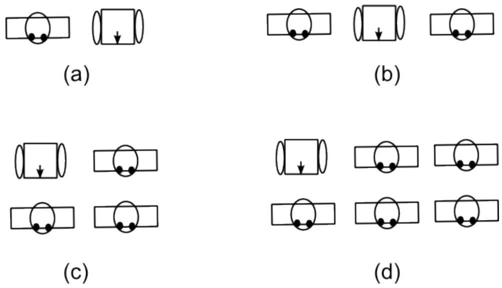


Figure 3.19: Group-robot size categories: (a) Group-robot Dyad; (b) Group-robot Triad; (c) 4 group-robot; (d) Large group-robot.

Table 3.7: Session times by locations

	Location											
Session	**1**	**1_2**	**2**	**2_3**	**3**	**3_4**	**4**	**4_5**	**5**	**5_6**	**6**	**Total**
S01	0:03:00	0:02:00	0:01:00	0:02:00	0:02:00	0:01:00	0:01:00	0:00:00	0:01:00	0:00:00	0:14:00	0:27:00
S02	0:02:00	0:03:00	0:02:00	0:01:00	0:03:00		0:00:00		0:00:00			0:11:00
S03	0:04:00	0:02:00	0:03:00	0:01:00	0:03:00		0:02:00		0:01:00	0:01:00	0:00:00	0:17:00
S04	0:02:00	0:03:00	0:03:00	0:03:00	0:03:00		0:02:00	0:00:00	0:01:00	0:00:00	0:01:00	0:18:00
S05	0:06:00	0:01:00	0:02:00	0:02:00	0:03:00	0:00:00	0:01:00	0:01:00	0:00:00	0:01:00	0:03:00	0:20:00
S06	0:02:00	0:02:00	0:02:00	0:02:00	0:04:00		0:01:00	0:01:00	0:01:00	0:01:00	0:00:00	0:16:00
S07	0:17:00	0:04:00	0:01:00	0:04:00	0:02:00	0:02:00	0:02:00	0:00:00	0:02:00	0:00:00	0:00:00	0:34:00
S08	0:02:00	0:03:00	0:02:00	0:05:00	0:01:00	0:01:00	0:02:00	0:01:00	0:01:00	0:00:00	0:00:00	0:18:00
S09	0:02:00	0:07:00	0:04:00	0:01:00	0:05:00		0:07:00		0:00:00			0:26:00
S10	0:03:00	0:02:00	0:02:00	0:00:00	0:05:00	0:03:00	0:01:00	0:00:00	0:00:00			0:16:00
S11	0:03:00	0:01:00	0:03:00	0:01:00	0:02:00	0:01:00	0:02:00	0:02:00	0:03:00			0:18:00
Mean	0:04:11	0:02:44	0:02:16	0:02:00	0:03:00	0:01:20	0:01:55	0:00:37	0:00:55	0:00:26	0:02:34	0:20:05
StDev	0:04:25	0:01:41	0:00:54	0:01:29	0:01:16	0:01:02	0:01:49	0:00:45	0:00:57	0:00:32	0:05:10	0:06:26

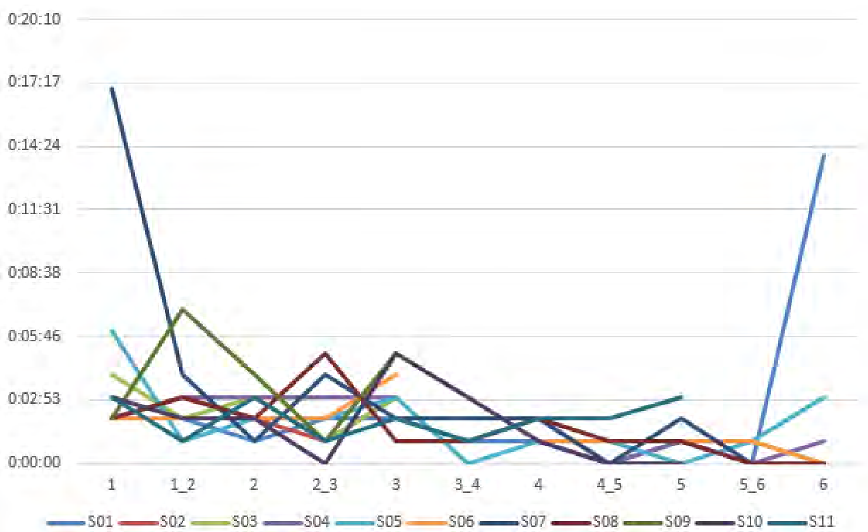


Figure 3.20: Graphic of session times by locations

each point correspond to the different sizes of groups detected in the sessions in each location. It should also be noted that the size of the groups can vary throughout the session, and even in one point, *i.e.* that people may arrive or leave at any point.

In location 1 a large percentage of L-groups (92 %) can be observed, possibly because it is the starting point of the robot and generates great expectation. As the guidance begins (location 1_2) we see that the percentage of L-groups decreases to 80% while other groups appear (triads 13% and 4HR groups 7%), This variation may be due because some of the people did not continue with the interaction or due to the physical limitations of the space (narrow corridor). Unlike 1, in location 2 a greater variation of the size groups can be observed (dyads 15 %, triads 8 %, 4HR groups 8 % and Large groups 69 %). Location 2 is a zone of little traffic in general, so it is presumed that people who were in 2 came mainly from point 1 next to the robot and at this point they are more dispersed. Likewise, during transit 2_3 a variation is observed in the groups (dyad 15 %, 4HR groups 31 % and Large groups 54 %). From location 3 to 5, the prevalence of large groups can be observed again, this may be because this is a high traffic area, so new visitors may tend to join the robot's guidance. Finally in the zone 5_6 (dyads 10 %, triads 10 %, 4HR groups 20 % and large groups 60 %) and at the point of arrival 6 (4HR groups 40 % and large groups 60 %) a variation is again seen in the groups. One possible explanation is that having interacted with the robot at points 3, 4 and 5, at the end, at point 6, people no longer pay attention from the robot.

Table 3.8: Relative frequency of group size by location

	Location										
Group Size	**1**	**1_2**	**2**	**2_3**	**3**	**3_4**	**4**	**4_5**	**5**	**5_6**	**6**
Dyad	0,00	0,00	0,15	0,15	0,00	0,00	0,00	0,00	0,00	0,10	0,00
Triad	0,08	0,13	0,08	0,00	0,00	0,00	0,09	0,00	0,00	0,10	0,00
Four	0,00	0,07	0,08	0,31	0,00	0,00	0,00	0,00	0,00	0,20	0,40
Large	0,92	0,80	0,69	0,54	1,00	1,00	0,91	1,00	1,00	0,60	0,60
Total groups observed	12	15	13	13	12	5	11	7	10	10	5

Spatial arrangements

From the 113 spatial arrangements observed, 52% corresponds to performer-audience formation, 29% goes to leader-follower arrangement, 7% have a river-like formation, 4%

Table 3.9: Relative frequency of spatial arrangements by location

	Location										
Spatial Arrangements	**1**	**1_2**	**2**	**2_3**	**3**	**3_4**	**4**	**4_5**	**5**	**5_6**	**6**
Vis-a-vis	0,00	0,00	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00
Circular	0,00	0,00	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,10	0,00
Cluster	0,08	0,00	0,00	0,00	0,08	0,00	0,09	0,14	0,00	0,00	0,00
Performer - audience	0,92	0,00	0,77	0,08	0,83	0,20	0,82	0,00	0,90	0,30	1,00
Leader-follower	0,00	0,67	0,08	0,62	0,08	0,60	0,00	0,86	0,00	0,40	0,00
River-like	0,00	0,20	0,00	0,08	0,00	0,20	0,09	0,00	0,10	0,10	0,00
Line-abreast	0,00	0,07	0,00	0,15	0,00	0,00	0,00	0,00	0,00	0,00	0,00
“<”-like	0,00	0,07	0,00	0,08	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Total observed	12	15	13	13	12	5	11	7	10	10	5

and 3% corresponds to cluster and line-abreast arrangements respectively and 2% goes to “<”-like, circular form and vis-a-vis formations (see Figure 3.22).

Table 3.9 summarizes the relative frequency of spatial arrangement observed at different locations. The performer-audience arrangement was observed most frequently at all robot stops, *i.e.* at face-to-face interactions. Also, in high traffic zones, a greater prevalence of the leader-follower formations could be observed *i.e.* in walking interactions. Also, two transit zones with the highest percentage of river-like formations (20 %), which are zones 1_2 and 3_4, should be highlighted. In zone 1_2 this formation could occur due to limitation in the physical space, which generated a clumping of the people around the robot, losing the leader formation. Since 3_4 is an entrance to a high traffic area, new individuals could join the interaction forming a crowd. The cluster spatial arrangement can be seen in the area 4_5 with a frequency of 14% due to its high traffic were crowds were formed. Finally, it is worth noting the several arrangements in zone 5_6 (40 % leader-follower, 30 % leader-follower, 10 % river-like, 10 % circular and 10 % vis-a-vis), once the robot passed through the high transit points 3, 4 and 5, and heading towards the final point with less agglomeration.

3.3.4 Discussion

Although the average time per session was 20 minutes, and in principle you could have up to 5 sessions per hour, ie 40 sessions a day, there were several factors that extended the time between sessions. For example, the dead times at the beginning and end of each session, either because there were no people interacting with the robot at the beginning, or that there were people interacting with the robot at the end; besides that the robot also had to cross all the way back to the place of departure. It would have been interesting to have a record of the number of visitors and the percentage of people who participated in the tour. However, it is difficult to keep track of how many people are at a certain moment in the science museum, since a variety of people come and go constantly, and the site is very large.

At the beginning of the study, it was set up that an observer would take pictures from 11 different points, as it can be seen in Figure 3.17. Due to the dynamics of the interaction, however, it was difficult to take the pictures at this points. For this reason we opted to reduce the takings of photos to the more significant observation points.

There are certain cases in which there exists a large number of people around the robot, however many are far away from it (see Figure 3.23). In these scenarios, there is the problem of defining which are the individuals that are part of the group. In terms of spatial arrangements, the people density around the robot may denote a social cohesion, and therefore the formation of a human-robot group. So it is proposed to use this metric in future works in order to define the human-robot groups more consistently. In this case, a high density in the human-robot group would define the area of group occupation, although a distance in which a social interaction can be carried out should also be considered.

Spatial arrangements were very dynamic in this type of interactions, so there were occasions when it was difficult to differentiate some kind of previously defined arrangements. This is happening often in transitions that involve changes in displacement, when the group goes from one motion state (*e.g.* static) to another (*e.g.* walking). In this situation, there is a space/time of negotiation where individuals adapt to the new state (from leader performer in the case of the previous example). In this analysis a particular arrangement has been observed, the “making the corridor” behavior, when the visitors make a corridor to the robot (see Figure 3.24). This arrangement has been identified when the Human-Robot group starts from a face-to-face interaction to a walking behavior. In this transition the robot starts to move against the direction of the individuals that are around around them, breaking their O space, so the individuals begins to make a corridor for the robot to continue its route. After the robot goes through this corridor the individuals continue accompanying the robot in its guiding.

3.3.5 Conclusions

A second study on group-robot interaction was carried out during a weekend at CosmoCaixa science museum. In the study, 11 sessions were observed for a total of 221 minutes. The -5 plant of CosmoCaixa resulted in an interesting space to carry out this type of study, since it allows the observer to have a privileged view of the interactions (from the plant -4) and at the same time generates a minimum of reactivity in people.

In this study the group-robot were characterized according to their size and their spatial arrangement during the guiding process. Observational method where used to describe the group-robot interaction.

In this study, the largest percentage of the group-robot observed were the larger groups (greater than 4 individuals), while the lowest percentage were dyads. This can be explained by the profile of the visitors who attend the science museum, which is presumed to be mostly accompanied by other people (*e.g.* friends, family or students).

Likewise, the observed spatial formations corresponded to the sizes of the groups, with the performer-audience and leader-follower arrangements being the ones that were found more frequently, and less frequently the “<”-type, circular and vis-a-vis arrangements.

Throughout the guided tour we could observe a great dynamic in the size and arrangement of the group-robot. In most cases, the groups behaved in a manner consistent with the established zones (zones for face-to-face interaction and for walking behavior). However, certain physical characteristics of the environment (such as narrow corridors) as well as the degree of occupancy of the visitors (high level of traffic) in certain areas may tend to modify the behavior of the groups.

There are situations where it was difficult to identify the groups, either in their size or in their form, as is the case where there are crowds or in the transitions from a face-to-face interaction to a walking behavior and vice versa.

Finally, during the transition from face-to-face to walking, we observed a particular type of spatial arrangement that we call “making the corridor”, where the robot breaks the O-space and the individuals must make a space for it to pass. In the design of the HRI, this type of arrangement should be taken into account, since this situation can inevitably occur, especially in crowded environments, so that in case of occurrence, this formation can be interpreted and executed in a fluid way by the users, in order to avoid blockages to the passage of the robot.

In the next section we will describe another REEM robot experience in CosmoCaixa, but which differs in terms of objectives, scenario, task and data analysis. In this case the robot will be a teacher assistant robot who accompany children during an educational robotic workshops.

3.4 REEM as a Teacher Assistant Robot

In this section we describe a third study with the REEM robot with a social service, but unlike the previous studies, the robot acts as a teacher assistant in an educational robotics workshop. This activity was developed at the end of November 2013 in a tech educational class at CosmoCaixa.

The present work aims at gaining understanding about the degree in which children feel the company of the robot as a social entity, through its "social presence" and its identity. These concepts were defined in Section 2.1.

Several groups of children interact with the robot in an edutainment (educational entertainment) activity. In comparison with the previous observational experiments, where the individual and group behaviors could be studied through direct observation methods, in this case self-report techniques were more pertinent to describe the children's perceptions and judgements about the robot and the situation. For this reason, a questionnaire was used as evaluation instrument.

Below, the study design of this experience is presented in Section 3.4.1. Next, the data analysis is described in Section 3.4.2. Results from the analyzed questionnaires are presented in Section 3.4.3. Finally, discussion and conclusions about the results obtained are presented in Sections 3.4.4 and 3.4.5 respectively.

3.4.1 Study Design

Objective

The objective of this study is to analyze the perception of social presence and identity of the robot in a group of children involved in an edutainment activity.

Scenario and Setup

The study was carried out on November 27, 2013 in one of the rooms of the museum (as shown in Figures 3.25a- 3.25b), within a program of educational robotics workshops. In order to favor the HRI, in this experiment the Wizard-of-Oz technique was applied. Hence, the robot was in part tele-operated, and its operator was outside the classroom and out of sight of children (see Figure 3.25c).

Task

Three roles were performed by the robot in the workshops: as a presenter, as a teacher's assistant and as an entertainer. As a presenter, the robot makes an introduction about the world of robotics and about itself. As an assistant, the robot moves around the classroom and attends to some children's questions. As an animator, the robot interacts with the children, greets them and takes a picture with them.

Procedure

The workshop program consisted of three main parts: an introductory talk, activities with different types of educational robots (for example bugs, Mindstorms NXT and Ugobe's PLEO) and the closing of the workshop. The duration of the workshop is 1 hour and it is expected to have a maximum of 5 workshops in the day.

At the beginning of the workshop, the REEM robot, with a female voice, was placed in front of the classroom. During the introductory talk the teacher gave way to REEM to talk about their skills (Figure 3.26a). Then while the different activities were being developed the REEM robot moved around the classroom with the objective of interacting and resolving some participants' concerns(Figures 3.26b– 3.26d). At the end of the workshop there was a space of time for the participants to interact freely with REEM (Figure 3.26e). Finally the questionnaires were delivered to be filled by the participants (Figure 3.26f.

3.4.2 Data Gathering

A questionnaire was used as the instrument to measure the perception of the social presence of the REEM robot. The questionnaire was composed of four sentences with a 5-points Likert scale of emoticons in order to measure the degree of social presence of the robot (Q1–Q4 in Table 3.10); and three questions related to the identity of the REEM robot (Q5–Q7 in Table 3.10). An example of the questionnaire can be seen in Apendix A).

Table 3.10: REEM social presence and identity questionnaire.

Q1	When interacting with REEM, I felt as if I did it with a real person
Q2	Sometimes I felt like REEM was really looking at me
Q3	I can imagine the robot (REEM) as a living being
Q4	Sometimes it seemed as if the robot had feelings
Q5	Give the REEM robot a name
Q6	How old do you think this may have
Q7	You think it's a boy or a girl

In order to establish the reliability of the social presence test we use the criteria of Cronbach's alpha [49]: Excellent (alpha>0.9), Good (0.7<alpha<0.9), Acceptable (0.6<alpha<0.7), Poor (0.5<alpha<0.6), Unacceptable (alpha<0.5).

The names that the children gave to the robot in Q5 were classified as human and non-human. Adicionalmente los nombres no-humanos pueden ser clasificados en dos tipos: mecanoide y mascotas. The ages assigned to the REEM robot in Q6 were classified by the following age groups: child, adult, elderly and non-human. El criterio para clasificar el nombre dependera del observador.

3.4.3 Results

A total of 124 children from the Montserrat School (Barcelona) with an average age of 11 years old participate in the workshops. Three questionnaires were discarded due to inconsistencies in their reading. From the total of children, a 55% percent were boys while 45% were girls.

Social Presence

The Cronbach's alpha calculated for the items about social presence of the robot was 0.60 for the five items. Hence, it can be established that Social Presence is an acceptable construct in this study.

In general, positive evaluations were obtained on the social presence of the robot. The question that obtained the best evaluation was Q2 (mean = 4,7; STDEV = 0,7), whereas the question with the smallest evaluation was Q3 (mean = 3,5; STDEV = 1,2) 3.27. The reason for Q2 to have a good rating may be because the question highlights the fact that the robot had the ability to look at users.

Identity of the robot

Identity of the robot is explored thorough the name assignment and the perception of age and gender.

In general, most of the names (63%) assigned to the robot by the participants were non-human (*e.g.* Ropmot, Robot, P-Bot, Smarkey, Runcky, Robit, Robi, Bret, Mucky) compared to human names (37%) (*e.g.* Rosa, Ramona, Miria, Rob, Josep, Monic, Nuria, Quica, Remesme).

On the other hand, in general it was observed that most of the robot age assigned from the children were in the group of children and adults, 44% and 47% respectively; followed by the non-human age (6%), the age of the older adult (2%) and the year of robot manufacture (1%).

It should be noted that non-human names have the highest number of non-valid ages (10.5 %) (*e.g.* 3022, 1494, 1E + 13, 1457) than with human names (2%)(*e.g.* 100 years old).

In the case of the gender, in general 65% of the participants indicated that the robot have female gender, while 31 % indicated that the robot have male gender. Only 3% indicated that the robot had no gender. In the same line, an interesting observation was that 5 participants attributed the gender of REEM by their voice (they wrote down this info in the questionnaire), 3 of them mention that robot is male while 2 mention that robot is female. It is also interesting to note that 3 participants indicated that the robot is “hermaphrodite” and one participant indicated that it was asexual.

3.4.4 Discussion

A fluid interaction usually requires short response times between the robot and the user. However, and due to the architecture used, the times between the operator and the robot were considerable, therefore the response times of the robot, and to a certain extent the HRI were affected.

Although the Cronbach's alpha reliability value of the social presence was on the edge of what is acceptable, the results generally show a perception of the robot's social presence in this experience above average.

The greatest amount of HRI could be observed in the role of animator at the end of each workshop. This is not the case for the assistant role. Therefore, it should be established in a clearer way the tasks performed by the robot and that improve the HRI. For example, within the design of the workshop activities you can include problems in which the robot helps with a clue for its solution.

Despite the fact that REEM robot head has a rather neutral appearance, his female voice apparently played a large part of the decisions to indicate that REEM was female. It's interesting to see that nevertheless, 31% answered the robot gender is male. One possible explanation is that the appearance of the robot's body has a rather masculine appearance.

3.4.5 Conclusion

A descriptive study on the perception of the social presence and identity of the REEM robot in a robotic workshop with children was carried out.

Despite the good results, we can mention that the perceptions of the social presence in this study could have better results if during the interaction the robot presented a more useful service for the participants.

In the design of the HRI, the tasks of the robot must be specified more clearly; likewise to encourage HRI, questions can be incorporated into the activities in which the robot can give some clue. In addition, it must be taken into account that the delay times between the operator and the robot can affect the HRI.

There are some aspects observed about the identity of the robot. One of them is that most children assigned a non-human name. The age that the children assigned to the robot corresponded mostly to the age of a child or an adult. It should be noted that the greatest inconsistency in the assigned age (*i.e.* non-human ages) corresponded to non-human names. The gender assigned to the robot was mostly female, possibly because of the voice it had.

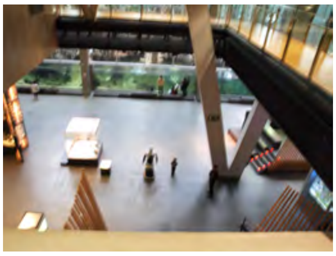
It would be necessary to explore in more detail the relationship between the factors of the name, age and gender assigned to the robot, and how this may affect the treatment or interaction between the human and the robot.

Until now, several HRI studies have been carried out in the CosmoCaixa science museum, during which the latest-generation robot REEM has played the roles of museum guide and assistant professor. As a result, several insights have been obtained about the behavior of the group-robot as well as the social presence of the robot.

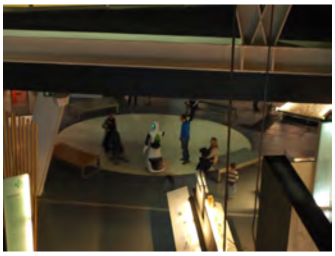
Because this study deals with the study of HRI in public spaces, we see the need to have more experiences of social interaction in other scenarios. In addition, the presence of the robot is required to be more "transparent" during the interaction, that is, the development of the robot is more focused on the service it provides than on its technical characteristics.

Therefore, we can see the need to develop a robotic platform that, although it is a certain part with limited technical characteristics, is evolving according to the social services it will provide.

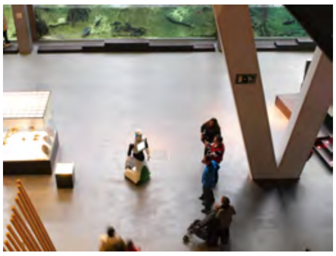
Following is the development of the Multipurpose Assistant Robot for the Study of Social Human-Robot Interaction - MASHI, based on several cycles of social interaction and the construction of the robotic platform.



(a)



(b)



(c)

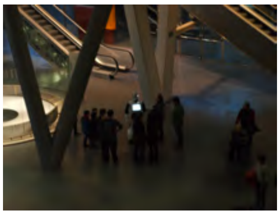
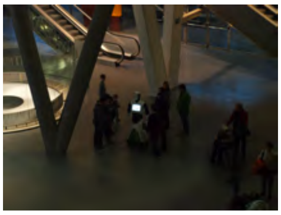
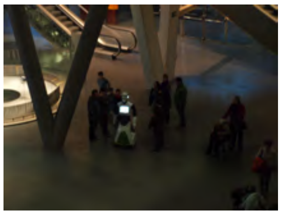
(d)

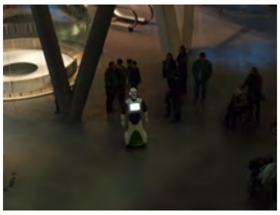
Figure 3.21: Group-robot sizes observed during the HRI: (a) Dyad; (b) Triad; (c) 4 Group-robot; (d) Large Group-robot.

(a) (b) (c) (d) (e) (f) (g) (h)

Figure 3.22: Group sizes observed during experiment 2: (a) performer-audience; (b) leader-follower; (c) river-like; (d) cluster; (e) line-abreast; (f) “<”-like; (g) circular form; (h) vis-a-vis.

Figure 3.23: Who makes up the human-robot group?


(a)

(b)

(c)


(d)

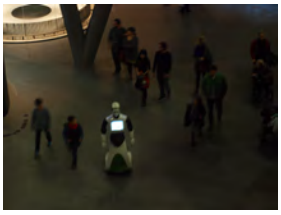
(e)

(f)

Figure 3.24: Spatial arrangements transitions: "making the corridor" behavior.

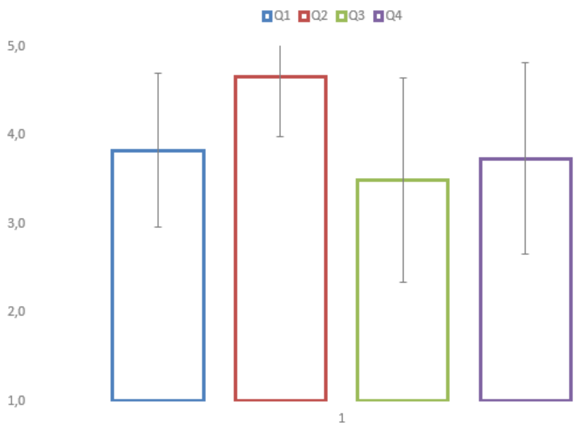

(a)

(b)

(c)

Figure 3.25: Scenario and setup for the teacher assistant robot experience: (a)-(b) educational robotics workshop scenario; (c) Operator for the Wizard-of-Oz HRI.

Figure 3.26: A robot teacher assistant session example: (a) Introduction; (b)–(d) HRI during the activities; (e) HRI during the closing of the workshop; (f) Participants filling out the questionnaire.



(a)

Figure 3.27: Social presence of REEM robot results.

Chapter 4

Developing MASHI

Once carried out the starting experiences with the sophisticated PAL Robotics' REEM robot, one of our questions when studying human robot interactions was to know to what extent the complexity of a robot favors or not the social interaction as well as the quality of the HRI service. Besides, there was not too much availability in the use of the REEM robot to study HRI for long periods in real environments.

In this chapter we describe the incremental processes of construction and programming of a robotic platform for the study of human-robot interaction in field experiments. Starting from a basic structure and programming, the robotic platform evolved in sequential sprints based on the services required and the experiences of interaction with users. To shorten the construction times of the robot, rapid prototyping techniques such as 3D printing were used.

In order to represent the main components of the robotic platform, a general scheme is shown in Figure 4.1. In this, you can differentiate two main components, the operator and the robot. The operator is a person in charge of controlling all the actions of the robot through a graphical user interface - GUI (operator's GUI). These actions are transmitted through a wireless network to the robot, which generates the HRI. On the robot side there is also a GUI (robot's GUI) which is managed by an expert to set the initial parameters of the platform.

During the construction process of the robotic platform several people were involved from different areas of specialization, such as industrial design, mechanical design, programming, among others. The robotic platform was designed and built entirely at the Universitat Politècnica de Catalunya · BarcelonaTech.

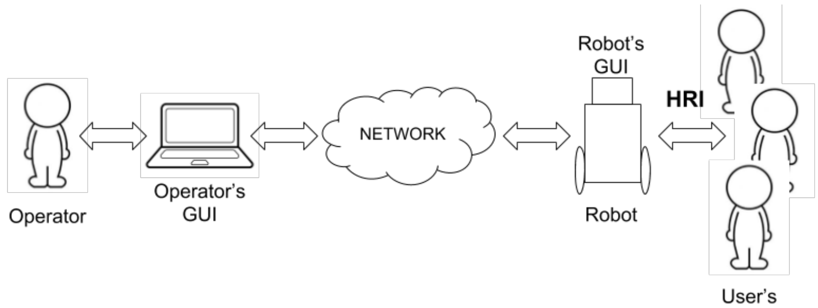


Figure 4.1: General scheme of the robotic platform.

This chapter is distributed in the following sections: in Section 4.1 the initial robotic platform is described. Next, in Section 4.2 the first development of our robotic platform is described as well as its validation in two outdoor trade fairs. Then, a second version of the platform and its validation in a technology fair is explained in Section 4.4. The development of a third version and its evaluation are exposed in Section 4.6. The development of the fourth version of our robotic platform, the robot arms, is finally presented in Section 4.8.

4.1 The Initial Robotic Platform

Below we detail the main characteristics of the initial telepresence robot, named Dave-Bot, used as a basis for iterative improvements.

4.1.1 Technical Description

Mechanical Structure and Appearance

The initial robot that was received is a mobile base with differential structure, *i.e.* with two drive wheels on the sides. A caster wheel on the back of the base and a stop as a replacement for the caster wheel on the front complete the differential drive locomotion. In the center of the base is placed a 660mm high mast; at its upper is placed a square aluminum structure serving as support for a netbook. In this case, the robot's appearance is purely mechanical, as see in Figure reffigDaveBot). Inicalmente este robot era utilizado como robot de telepresencia.

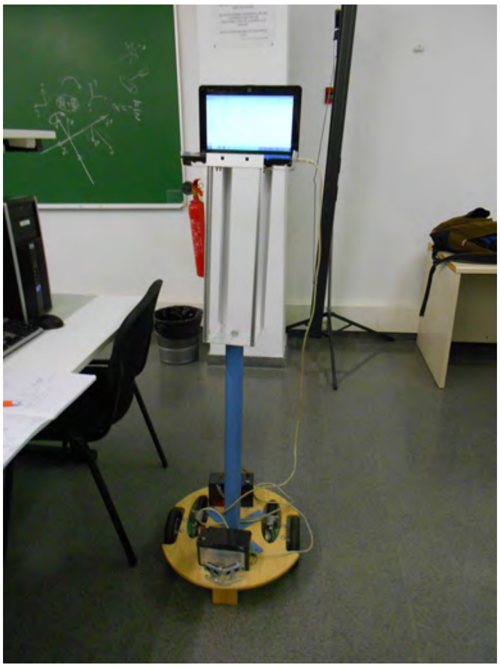


Figure 4.2: The initial robot.

Sensors, Actuators and Microcontroller Board

The robot is endowed with two ultrasonic sensors located on the front of the base as well as two encoders coupled to the motor axes (Parallax Position Controller Kit).

As actuators two worm-drive DC gear-motors are mounted on precision machined solid aluminum motor mount blocks. They can carry up to $27Kg$ of payload (Aluminum Drive System). Two H-bridge type motor controllers can handle a load of up to $25A$ in direct current (Position Controller Kit).

Listed sensors and actuators are connected to an Arduino UNO microcontroller. The Arduino board is connected to the netbook by using a USB cable. This board works as an intermediary between the computer and the actuators and sensors.

Power Supply

The robot is powered by two 12V sealed acid batteries (SLA), one of them feeding the motors and the other one the Arduino board. The netbook is powered by its own battery.

Information and Communication Technology

An *Asus Eee* netbook provides most of the computational and communication skills, while the screen serves to show the operator's face for telepresence tasks.

The *vLine* platform have been chosen as the communication system in order to generate a cloud infrastructure for the *WebRTC* API. In this form we provide our browser with Real-Time Communications (RTC) capabilities via simple API, avoiding to be worried about all the potential technical problems. Having several sets of APIs, WebRTC can be used as an out of the box solution for making online video chat, as well as implementing behaviours with custom code. What is perhaps even more important, both approaches can be mixed to only implement those elements needed for custom solutions. Important issues solved by vLine include: establishing connections, NAT traversal, signalling, user authentication, and contact management.

A server in both, operator and robot sides is necessary to handle those tasks. In this case the standard code provided by vLine was used, then an app was developed on top of it. Since the solution is based on webRTC, it is portable, plugin free, and video streaming is performed in a peer-to-peer, browser-to-browser mode.

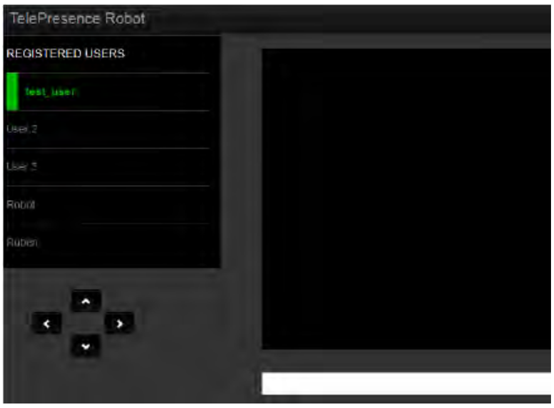
In order to provide communication between the browser and the robot control board (through the COM port) the plugin jUART has been used. It is compatible with all major browsers and operating systems . It provides a very simple API to send messages to the chosen COM port. Both, the operator side and the robot side work on graphical interfaces, as shown in Figure 4.3

4.1.2 Communicative Skills and Interactive Behavior

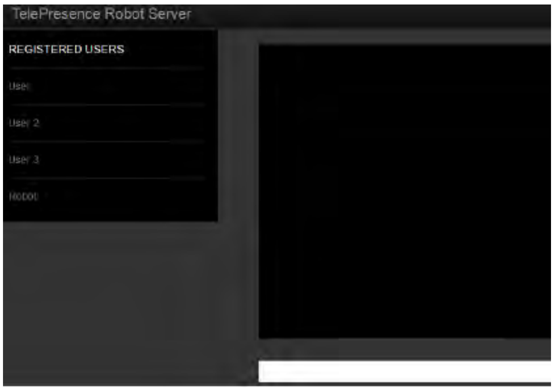
The initial robotic platform DaveBot is only able to navigate. It moves through the motion provided by the base, which is commanded by the operator through a graphical interface (see Figure 4.3a).

4.1.3 Discussion and Conclusion

The robotic starting platform has several advantages, such as: motors can carry up to $27Kg$ of payload, which can be attached to a structure in the mast of the center; Vline system allows to establish connections between the operator and the robot in a simple



(a)



(b)

Figure 4.3: Operator and Robot Graphical User Interface.

way, completed with a system of user authentication and contact management; allows transmission (full duplex) of video, audio and data in real time.

Despite these good features, we can list several limitations that prevent us from using it as a platform for the study of HRI in real environments, such as: the movements of the robot are very discrete so there is no softness at the moment of start or stop of the robot. From the mechanical point of view, the fact of carrying the laptop at the top makes the platform unstable by the moment of inertia that is generated. Moreover, the square structure was not well fixed with the mast, so it is more unstable and unsafe. Another security issue is that the devices were too exposed to the outside. The batteries also were not attached to the base and could generate safety problems. From the software point of view, both the operator side and the robot side required the execution of a server, hence being very resources consuming. This system also requires a registration and access to the Vline application on the Internet, limiting the possibilities in case there is no connection to it.

Once the benefits and limitations of the initial robotic platform have been analyzed, we set out to do a first cycle of development to cover the main requirements that our platform must accomplish for HRI in-the-wild.

4.2 MASHIv01: Robot Structure and Morphology

With the initial robotic platform already working, an analysis of the main problems to be solved was carried out. In this section we describe the first cycle of analysis and development of the robotic platform.

4.2.1 Analysis and Requirements

With the aim of designing and building a mobile robot to study HRI in open environments and based on the initial platform and the lessons learned in our previous field study, the following end-user requirements of the robotic platform were defined:

- Usually, the mobile robot has been used in indoor and on flat surfaces.
- Safer mobile robot, since the HRI will be carried out with a wide variety of people, including children or seniors. Therefore it is required to make use of light materials (*e.g.* cloth).
- Modular and scalable structure, which has the capacity to adapt new devices (*e.g.* depth cameras to detect people or to give navigation functionality in the robot).
- Adjustable height, in order to adapt the camera to the height of people.
- Appear more attractive and stylized by means of lightweight structures (*e.g.* 3D printing and fabric) to recreate more stylized volumes (*e.g.* umbrella structure).
- For the study of HRI based on spatial relationships it is necessary to have a wide field of vision around the robot. That is, the robot can recognize people around it at least at a social distance.
- Easily controlled by the operator (*e.g.* keyboard, mouse)

- Teleoperated mobile robot, with the possibility of being able to work as a teleoperated robot (*i.e.* a robot that represents a person who is physically in another location, but that users know that they are interacting remotely with this person) or as a telerobot (*i.e.* a robot that has its own identity, with a head, face, arms, which despite being controlled remotely, no longer represents the person behind the controls (operator), but represents a robotic agent.)

- Real time communication, that allows a smooth HRI without delays. The communication range between the operator and the robot should be approximately $100m$.

- Better control of the base motion (*e.g.* motion in a straight line).

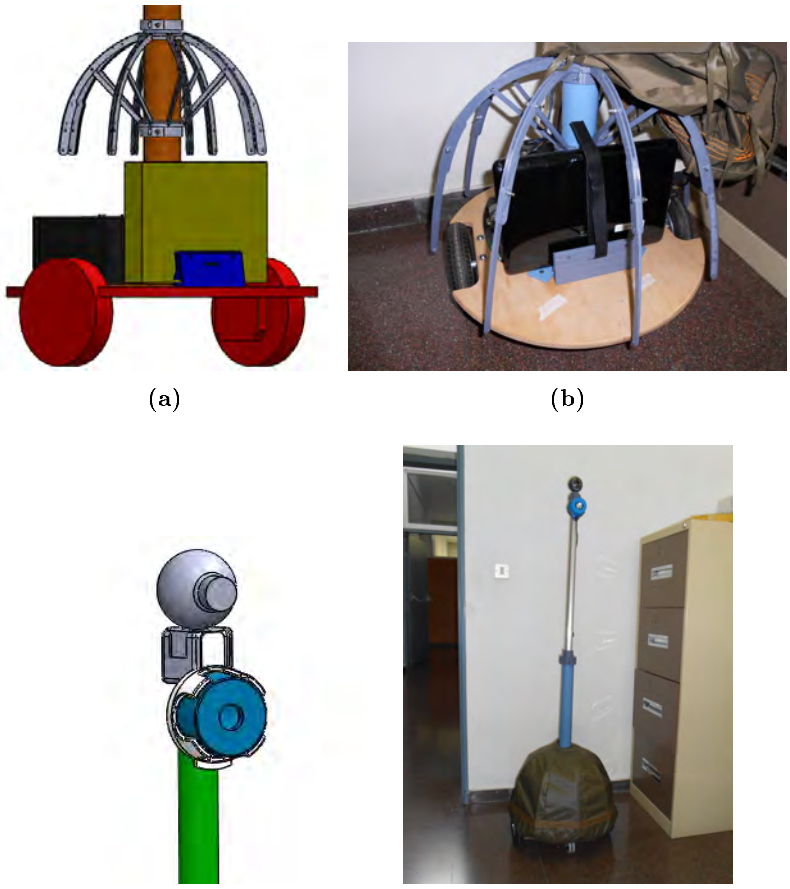
4.2.2 Implementation

Mechanical Structure and Appearance

In the first version of MASHI some sprints were completed in the development of robot's mechanical structure, giving a higher priority to safety than to form.

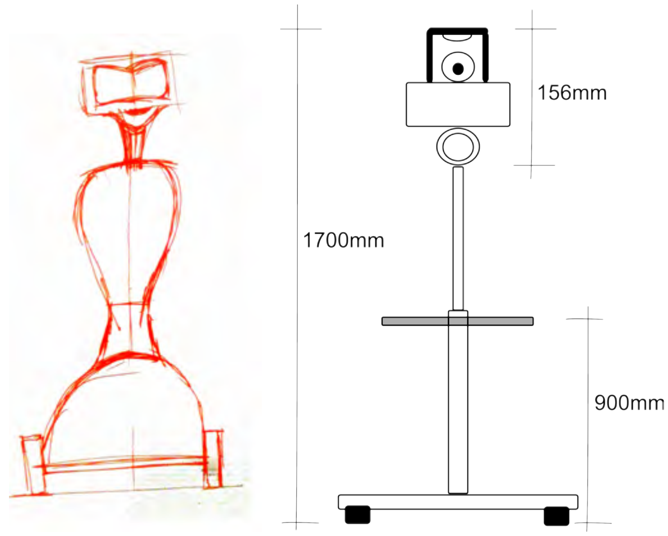
First sprint. In this first sprint goal was focused on lowering the greater amount of the weight of the robot towards the base, with the purpose of reducing the moment of inertia exerted in the upper part of the robot. For that reason the netbook was placed in the base. A special part was 3D printed to fix it. The battery was also fixed in the base (Figure 4.4a). An aluminum tube was placed inside the mast, which allowed a webcam and a portable speaker to be located at the top, as shown in Figure 4.4c. To cover the base of the robot, an umbrella-type structure covered with a cloth was used (see Figure 4.4a). To maintain a better stability of the mobile base, a caster wheel was placed on the front of the base. The different elements of both the fixation and the structure of the base cover were designed and manufactured using 3D printing techniques.

Second sprint. In this second sprint we address the following topics:



(a) (b) (c) (d)

Figure 4.4: MASHI v01 first sprint: (a) Bottom structure model; (b) Bottom structure; (c) Upper structure model; (d) MASHI v01 first sprint robot.



(a) (b)

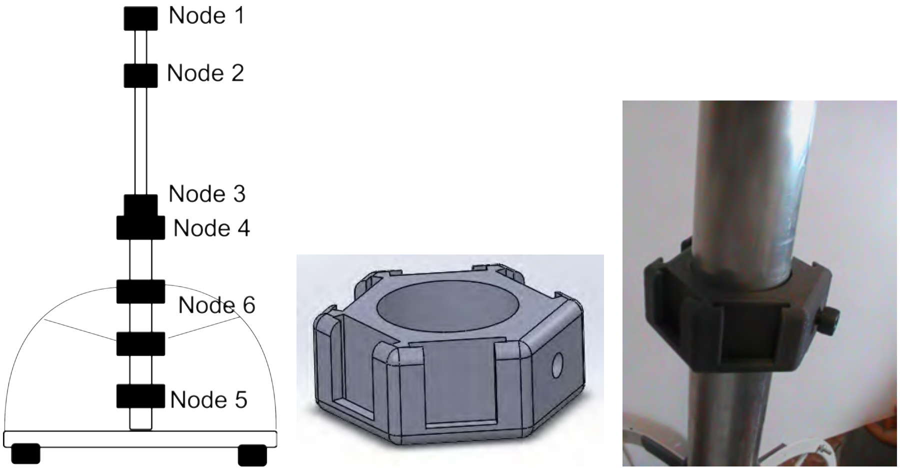
Figure 4.5: MASHI v01 second sprint: (a) Robot sketch; (c) Robot structure design.

Design and shape of the robot. Although it has a robot design like the one in the Figure 4.5a, in this sprint emphasis has been given to the functionality and proportions of the robot.. To show the operator's face, a display was placed on top of the robot. In addition, a webcam was placed above the display for better maneuverability of the operator and field of vision. Additionally, the web camera has a degree of freedom, either to observe the scene in front of the robot or to see the scene around the robot through an omnidirectional mirror, as shown in the Figure 4.6c.

Modularity. The concept of nodes was developed, wherein different types of devices can be hooked to the central mast (Figure 4.6).

Camera and mirror mechanism. In order that the robot can have a wide field of vision and thus recognize the people around it, a camera-mirror mechanism is proposed, in such a way that by means of a single camera one can have both a short field of vision to see things right in front of the robot as well as a mirror system to see around it, as shown in Figure 4.7.

The different development sprints in the mechanical structure of MASHIv01 can be observed in Figure 4.8.



(a) (b) (c)

Figure 4.6: MASHI v01 modularity through the concept of nodes: (a) Nodes distribution; (b) Node 2 model; (c) Node 2 printed.

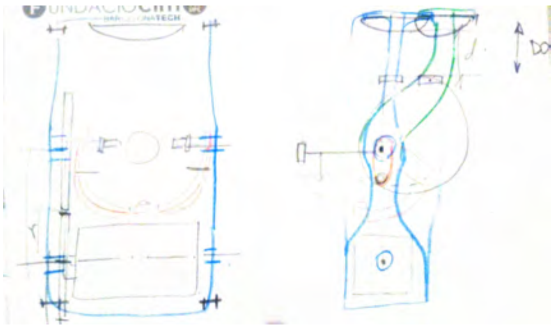


Figure 4.7: MASHI v01 camera and mirror mechanism sketch.

Sensors, Actuators and Controllers

The webcam and the built-in microphone were used as sensors for the human operator to see and hear the environment.

In MASHIv01 a Dynamixel AX-12A servomotor was incorporated to allow the vertical movement (tilt) of the webcam. The OpenCM9.04 board was used to control the servomotor. The OpenCM unit will receive commands (slave mode) of movement through the Arduino board (master mode) through I2C communication.

The velocity control of the base is performed through discrete movement commands such as forward, backward, left and right. The velocity control is in open loop. To reduce the abrupt change between movement commands (for example when the robot

(a) (b) (c)

Figure 4.8: MASHI v01 structure evolution: (a) Sprint 1; (b)-(c) Sprint 2.

is still and a forward command is sent) an algorithm was implemented to smooth the transition between movement commands.

Information and Communication Technology

The communication platform WebRTC was used, since it allows the transmission of audio, video and data in real time, which are part of the requirements of our robotic platform.

The implementation has been built on top of an already existing architecture. The already existing implementation is built with NodeJS[1], which is a cross-platform runtime environment allowing for easy creation of server-side and network applications.

NodeJS builts on the Google Chrome JavaScript API[2], for which the code is composed of JavaScript, HTML and CSS.

The system architecture is subdivided into an *operator part*, which runs on the operator's computer and provides the control interface for MASHI, and a *robot part*

[1] NodeJS: `http://nodejs.org/`
[2] Google Chrome JavaScript: `https://developer.chrome.com/extensions/api_other`

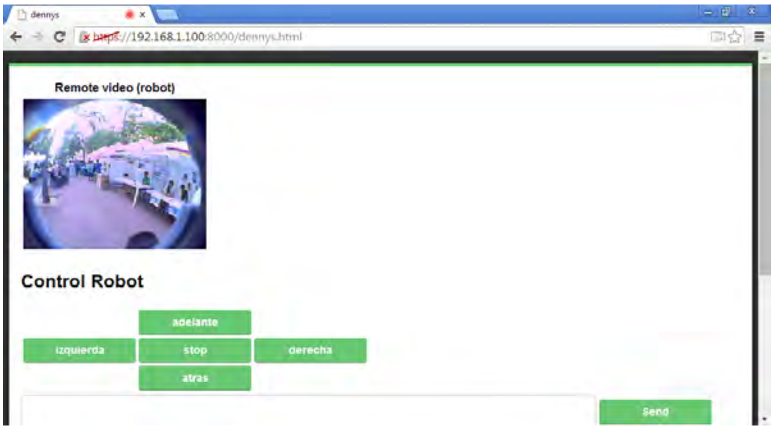


Figure 4.9: MASHI v01 operator interface.

which runs on MASHI, receives the operator's commands and executes them.

On the robot server, the communication between the robot's browser and the operator as well as the communication between the robot's computer and the Arduino were implemented.

In the browsers of the robot and of the operator, the functions that allowed the transmission of data from the operator to the robot to perform the movement of the base and the text-to-speech were implemented.

The graphic interface of the operator is divided in three sections: a view of the front camera of the robot, the buttons to operate the movement of the base and a text box to perform the text-to-speech function, as seen in Figure 4.9.

Communicative Skills

The transmission/reception of the audio between the operator and the users allows a verbal communication. The operator's video allows non-verbal communication through their facial expressions.

The movement of the robot with a certain cadence in its movement, varying the speed of advance of the robot, allows to replicate the cadence that humans have when they walk, giving a more social impression to the act of walking of the robot.

4.2.3 Discussion and Conclusion

On the structure of the robot, the fact of lowering the weight of the computer (netbook), as well as replacing the square structure with a tube as continuation of the mast, allowed the moment of inertia to be considerably reduced. These modifications, coupled with a smoother transition between robot movement commands through programming, allowed the robot to be safer than in the previous version.

Although the robot remains still with a mechanical appearance, anthropomorphic proportions of the head were considered in the position of both, the display and the speaker (see Figure 4.6c).

The communication between the Arduino and the OpenCM boards through I2C protocol could not be solved, so it was decided that both the Arduino and the OpenCM would handle independent USB (COM) ports.

The mechanism of the omnidirectional mirror could not be implemented due to drawbacks in the dimensions of the mirror. We must look for other alternatives to have an omnidirectional image of the environment. In the front camera of the robot a fisheye type lens is attached so that the operator is empowered with a greater field of vision.

Although a social movement has been implemented for the platform, with the speed of the robot as a sinusoidal function, this funtion has not been validated to know the perceptions of people about the movement.

Developments in this first version, although quite basic, would allow us to take the platform to in-the-wild environments to observe how to test the functionalities and report reactions from the people. In the following section we describe the first HRI tests of our platform in crowded environments.

4.3 Validation of MASHIv01 in Trade Fairs

To validate the development of the MASHIv01 robotic platform, several experiences of social interaction were realized in open environments. Two experiences were carried out in the Municipality of L'Hospitalet de Llobregat - L'H in the framework of the Trade Fairs in the Districts of Pubilla Casas and La Florida. The robot was used as a means to attract the attention of people and to promote the edutainment activities of the CortoCircuito Robotics Club.

4.3.1 Scenario and Setup

The HRI experiences of MASHIv01 were carried out during the trade fairs during the festivities of the Pubilla Casas and La Florida neighborhoods. Tents were placed along a street so that the businesses of the neighborhood and the different social organizations could exhibit their products or services. Our robotic platform was used to promote out-of-school educational/entertainment services offered by the "Club de Robotica CortoCircuito".

Although there is no pre-established script, there was a behavior of the robot focused on maintaining the attention of the users. Thus, at the beginning the robot walks around the checkpoint in search of the interaction. When someone approaches or passes near the robot, the robot salutes, in which case if it has a positive response the robot tries to maintain a dialogue with the user(s) about the activities of the Club, until the people leave. Then the robot says goodbye and walks around again looking for new users.

To carry out the HRI, the robot has some features, such as the base motion, at the top (at the level of the head) is a 7 in screen to show the image of the operator. On the screen there is a motorized web camera so that the operator can have a wider view of the environment. At the bottom of the screen there is a speaker that reproduces the voice of the operator, as shown in Figure Figure 4.10. With these characteristics, the interactive behaviors are based mainly on the communication that the robot can establish through the voice of the operator, while the non-verbal communication is very limited, and basically falls on base robot motion.

4.3.2 Data Analysis

In this experience we want to observe the interactions of people with the robot. From a photo camera, images were used to analyze the experience, same that were taken by an external observer. The photos were taken whenever there was some kind of interaction between the users and the robot.

From them, we will describe the typologies of the groups that interacted with MASHI. An *ad hoc* coding scheme grouped by composition and size of the groups was established, as shown in Table 4.1.

To continue with the quantitative analysis, the group description data were tabulated taking into account the coding system in Table 4.1 applied to 94 HRI images

(a) (b)

Figure 4.10: MASHI v01 in Trade Fairs: (a) At "Pubilla Cases"; (b) At "La Florida".

Table 4.1: Encoding scheme for group description

Category	Encoding Scheme
Composition	Children (6-11 years old), youth (12-17 years old), adults, mixed groups
Size	Dyadic, Triad, 4 Group-robot, Larger group-robot

taken along the trade fairs. The identification of the groups in the images was performed through the interpretation of non-verbal language among their members such as body orientation, gesticulations and spatial cohesion of the group.

4.3.3 Results

The results obtained are shown in Table 4.2. Regarding the composition of the groups that interact with MASHI, 26.60% of the groups that interacted with the robot were only children, 9.57% were only young, 46.81% were only adults and 17.02% were mixed groups of children and adults.

Regarding the size of the groups, 47.87% of the individuals showed interaction with the robot individually, while 52.13% did so in groups: 37.23% of the groups were triads, 5.32% were 4 group-robot and 9.57% were large group-robot.

The analysis of the size of the group reveals that the individuals who interacted alone with the robot were 15.96% composed of children (Fig. 4.11d), 5.32% for young people and 26.6 % for adults (Fig. 4.11a). The triads were 7.45% children, 4.26% young, 15.96% adults (Fig. 4.11c) and 9.57% children and adults. 4 Group-robot were made up of 3.19% children, 1.06% adults and 1.06% children and adults. Large group-robot were formed by 3.19% of adults and 6.38 % of children and adults.

Table 4.2: Group description of HRI in the Trade Fairs.

	Age				
Group Size	Children	Young	Adult	Children + Adult	Total
Dyadic	15.96%	5.32%	26.60%	0.00%	47.87%
Triad	7.45%	4.26%	15.96%	9.57%	37.23%
4 Group-robot	3.19%	0.00%	1.06%	1.06%	5.32%
Large Group-robot	0.00%	0.00%	3.19%	6.38%	9.57%
Total	26.60%	9.57%	46.81%	17.02%	100.00%

4.3.4 Discussion

Due to the massive attendance of people at the fair and for safety reasons, the robot's motion was rather limited and not very fluid.

Given the massive attendance of people at fairs, it is necessary to be able to control the speed of the robot. In environments or situations with high density of people, the speed of the robot must be minimal for safety reasons.

The structure of the mobile base of the robot presented several drawbacks to cross different obstacles. Because the fair's surroundings were streets and sidewalks, it was very difficult for the robot to change between street and sidewalk, even through the access ramp for people with disabilities. For these cases a less rigid mobile platform is required, with damping in the wheels and with bigger driving wheels.

Even though the shape of the robot was quite simple, people did seem to distinguish the different parts of the robot because of the way they interacted with the robot. One of the elements that could influence the identification and integration of the robot was

(a) (b) (c) (d)

Figure 4.11: HRI experiences with MASHIv01 in Trade Fairs: (a) Side by side walk; (b) Interacting by voice; (c) Capturing attention of adult users; (d) Capturing attention of children.

the use of a shirt, which could give a more personal image to the robot. The robot was dressed differently in the two fairs, as shown in the Figure 4.11, due to a constant improvement in the appearance of the robot. It was observed that at least two children took the arm of the robot during the interaction at the first fair (robot wearing a T-shirt Figure 4.11d), but it can not be established in which way a relation of this with the results obtained.

While in general the surrounding noises were quite high, there was much variability in the level of outside noise during the day. Therefore, the operator must have the ability to control the volume of the robot's audio system.

From the point of view of the HRI, it was observed that in this context of these fairs the largest number of groups were individual (*i.e.* 1H-1R) and dyadic (*i.e.* 2H-1R), 47.87% and 37.23% ; and the ages were composed mostly of children and adults, 26.60

% and 46.81 % respectively. Unlike the previous study with the REEM robot, where larger and more varied groups could be observed, in this case the groups were rather small and with little participation of the young people. An explanation can be the environment itself, where youth participation could be scarce in this type of events in general. Among the factors that could have affected the size of the groups were possibly the type of task performed by the robot and the noise. An exhibition guide task can summon a greater number of people. In addition, an environment with a lot of outside noise can divert the attention of users who may be further away from the robot.

Due to technical problems it was not possible to show the operator's images on the robot's screen, nor to control the movement of the camera. In spite of that, it was possible to observe several interactive behaviors of people towards the robot, especially on behalf of the children. In an environment where a communication channel may be affected (*e.g.* voice communication due to loud noise), it is important to establish other communication channels (such as gesticulation through the head motion).

4.3.5 Conclusions and Future Work

An experimental robotic platform with a mobile base, anthropomorphic, modular and scalable structure for the study of human-robot interaction (MASHI) has been designed and built. A first exploratory study was conducted to analyze the types of individuals and groups that interacted with MASHI. In addition, as a result of these interaction experiences, important lessons have been learned about the design of the robotic platform.

There are major challenges in the mechanical and behavioral design of MASHI, in such a way that the robot can count with more interaction channels that improve both verbal and non-verbal communication. For example, adding an articulated head or arms to the robot opens up new possibilities to add interactive behaviors to the robot, such as gestures, gaze, etc; his also improves the anthropomorphic proportions of the robot.

The interaction experiments have been carried out outdoors and to increase mobility and accessibility in open and natural environments, the development of the assisted locomotion system of dynamic damping systems is planned in both the basic structure and the configuration of the wheels.

The following section describes the development of MASHIv02, especially in the development of the head of the robot. Several implications in the design of hardware and software had to be taken into account in this development.

4.4 MASHIv02: Robot Head

In this second version of the robot, and based on the HRI experiences carried out previously, several improvements have been performed. The requirements of this second version are described below. Then the different implementations in the robotic platform are described. The discussion and conclusions of this second version are exposed at the end of this section.

4.4.1 Analysis and Requirements

In the previous reported HRI experiences there were cases in which people approached the robot to hear some response. If the environment is quite noisy, it is very difficult listen to the robot, and the use of non-verbal language would be necessary. In interactions many times the robot must respond affirmatively, negatively or "I do not know". Therefore, in order to expand the communication modes between people and the robot, the development of an articulated robot head is proposed. On the one hand the head will allow non-verbal communication, but on the other hand it will allow the operator to expand his field of vision just as a person does.

4.4.2 Implementation

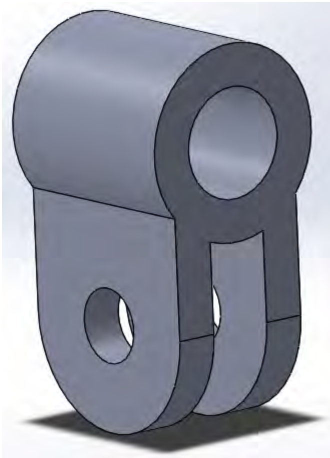
Mechanical Structure and Appearance

In order to give a volume to the body we use a structure similar to that of the cover of the base, that is to say of type umbrella. This volume gives the robot the appearance of a body, agency property, but with a fairly light weight (see Figure 4.12).

For the structural part we use the concept of connector design, that is, design and build 3D connectors to attach special devices (such as camera, display) and preferably make use of existing materials in the market (such as PVC rods) for the case of the structure of the robot body.

The articulation of the head has 3 DoF: pitch, yaw and roll, which allow the vertical, horizontal and inclination movements of the head (see Figure 4.12). The structure of

(a) Umbrella type upper body.



(b) 3D printed design for the umbrella connectors.

(c) Head of MASHIv02, internal view.

(d) Head of MASHIv02, cover.

Figure 4.12: MASHI v02 mechanical structure and appearance improvements.

the head must load the weight of the display and the camera mounted just on top of it.The speaker will be located at chest level.

Finally, a head housing was designed, both to protect the mechanical part of the head and to give it a better visual appearance (see Figure 4.12).

Sensors, Actuators and Controllers

The camera and the microphone of a webcam were used as sensors for the human operator to see and hear the environment.

MASHIv02 head has three Dynamixels AX-12A servomotors that were used to allow the pitch, yaw and roll movements. The OpenCM9.04 board was used to control the servomotors. The OpenCM will receive the head commands from the robot computer through a USB port. The commands of movement of the head are discrete values to make pitch up, pitch down, yaw left, yaw right, roll left, roll right and head at a central position.

The velocity control of the base is performed through discrete movement commands such as forward, backward, left and right. The speed control of the motors use a PID algorithm control, smoothing transitions between changes in the motor speeds.

Information and Communication Technologies

The communication platform WebRTC was used, since it allows the transmission of audio, video and data in real time, which are part of the requirements of our robotic platform. On the robot server, it allows the communication between the robot's browser and the operator browser. Also the communication between the robot's computer with the Arduino (robot base) and the OpenCM (robot head) were implemented. In the robot and operator browsers, the functions that allowed the transmission of the data motion commands (base and head motion's) from the operator to the robot and the text-to-speech were implemented. The graphic interface of the operator is divided in several parts: a view of the front camera of the robot, buttons to operate the base motion, buttons to operate the head motion and a text box to perform the text-to-speech function.

Communicative Skills

Both, the text-to-speech and the transmission/reception of the audio between the operator and the users allows a verbal communication. Besides, the operator's video allows non-verbal communication through their facial expressions. Moreover, another important mode in non-verbal communication is the movement of the head implemented, which will initially allow the robot to respond affirmatively (pitch movement), negative (yaw movement) and doubt (roll movement). The movement of the base also generates a type of proxemic behavior and spatial formations when there is an HRI experience.

4.4.3 Discussion and Conclusions

In the second version of the robot, the implementation of the head and its movements signified a great advance both, for the appearance of the robot and for non-verbal communication. However, due to safety issues of the mechanical structure of the head, movements of the head are very limited in the range of rotation and the speed of the servomotors. Improvements could be made by playing with these values to denote some type of behavior (for example, a faster and wider movement may denote a more open and extroverted behavior). In addition there are problems in the central position of the robot head, which is the position of the robot facing forward. This position is not a rest position of the robot, and given the current mechanical structure the servomotor produces a constant effort. For this reason, after a certain time of operation, the motor that performs the pitch movement is blocked by overheating. It will be necessary to look for new alternatives in the mechanical design to avoid this overexertion of the servomotor.

Another drawback detected has been the design of the head housing. Although in the design program the pieces fit perfectly, actually there is a difference with the dimensions of the pieces printed in 3D and these did not fit very well in their joints. Another point to discuss is that the current housing can not be assembled and disassembled easily, making it difficult to access the inside of the head.

In the following section we describe an experience of HRI held at the Smart City Expo World Congress in Barcelona with the MASHIv02 robotic platform.

4.5 MASHIv02 HRI Experience in SCEWC 2014

To validate the development of the MASHIv02, an HRI experience were realized in the Smart City Expo World Congress in Barcelona. The robot was used as a means to attract the attention of people to L'Hospitalet town hall stand.

The scenario and the setup of the MASHIv02 robotic platform for this interaction are described below. Next, in Subsection 4.5.2 the data analysis procedure is detailed, then in Subsection 4.5.3 the results about the HRI experience will be described. The conclusions and future works are mentioned at the end.

4.5.1 Scenario and Setup

The HRI experience was carried out at the Smart City Expo World Congress in Barcelona, during three days of the month of November 2014.

In order to attract the attention of visitors, the robot (through the operator as a wizard) will be moving around the stand of the Municipality of L'Hospitalet (as shown in Figure 4.13a). As in the HRI experience at the fairs, initially the robot moves around the stand and when someone approaches the robot, it waves and establishes an open dialogue until the user leaves, in which case the robot says goodbye and go back and forth in search of a new interaction. On this occasion the robot also has the movement of the head to be able to gesticulate and improve non-verbal communication.

In the operator interface, the control buttons of the movements of the base and the head were disabled and hidden, and in their absence they were controlled by means of a keyboard. In addition this allowed to have more space on the screen to see the image of the robot's camera (see Figure 4.13c).

4.5.2 Data Analysis

Like the previous HRI, in this experience we want to observe the interactions of people with this new robot version. From a photo camera, images were used to analyze the experience, same that were taken by an external observer. The photos were taken whenever there was some kind of interaction between the users and the robot.

From them, we will do a description of the groups that interacted with the robot based on an *ad hoc* coding scheme grouped by composition, size, F-formation and Proxemic distance, as shown in Table 4.3.

4.5.3 Results

The group-robot description were tabulated taking into account the coding system in Table 4.3 applied to 30 HRI images taken along the congress. The identification of the groups in the images was performed through the interpretation of non-verbal language among their members such as body orientation, gesticulations and spatial cohesion of the group.

Regarding the composition of the groups that interact with the robot, 96.00% of the groups that interacted with the robot were only adults , while only the 4.00% were

(a)

(b) (c)

Figure 4.13: MASHI v02 in SCEWC2014: (a) L'H stand; (b) Robot body; (c) Operator's GUI.

only children. As the event is more of an executive/technical nature, it was expected to have the type of adult user.

Regarding the size of the groups, we can indicate that the vast majority, a 85.18%, they were groups of dyadic type, while 7.41% was triads and also 4 group-robot. Again, the characteristic of the event could have influenced this result, since the attendees were largely executives and they could attend the event alone.

On the other hand, the spatial relationship reveals that, as might be expected, the largest number of groups were face-to-face type (62.96%), followed by side-by-side type formations (25.93%) and in smaller percentage were the circular formations (11.11%). The side-by-side formations corresponded mainly to people who took pictures with the robot. In terms of proxemic distances, most interactions occurred at a personal distance (51.85%), followed by a social distance (33.34%) and to a lesser extent an intimate distance (14.81%). It should be noted that the intimate distance was when the person

Table 4.3: Encoding scheme for group description

Category	Encoding Scheme
Composition	Children (6-11 years old), youth (12-17 years old), adults, mixed groups
Size	Dyadic, Triad, 4 Group-robot, Larger group-robot
F-formation	Face-to-face, side-by-side , circular form
Proxemic distance	intimate, personal, social, public

hugged the robot or while the person was taking a picture with it.

Among the interactive behaviors observed, 18.52% corresponded to people taking pictures to the robot (Figure 4.14a), and in the same percentage corresponded to people who took pictures with the robot (Figure 4.14b). A 3.7% the person expressly embraced the robot (Figure 4.14c), while a 7.41% hug the robot to take a picture (Figure 4.14d).

4.5.4 Discussion and Conclusion

During the Smart City Expo World Congress 2014 the Mashi V02 robot caught the attention of 16 visitors. There were technical problems that affect the number of HRI, and that did not allow for a higher control of the robotic platform. One of them was the control of the robot head; in a certain moment the movement was blocked. This bug, however, was reported. It was due to an increase in temperature in the servo that controlled the pitch since the servomotor that controlled this movement was always supporting the load of the head. Another problem added to the mechanical structure of the head was that the parts of the different joints of the head were very forced.

In the HRI analysis, it could be observed that, according to the rather formal and executive context of the event, the greatest number of people who interacted with the robot were adults. The groups formed were mainly dyadic, the proxemic distances of HRI were personal and social, while the type of formations were face-to-face and side-by-side. The type of side-by-side training occurred mainly when people took a picture

(a) (b) (c) (d)

Figure 4.14: Several interactive behaviors: (a) people taking pictures to the robot; (b) people taking pictures with the robot; (c) user hugging the robot; (d) user hugging the robot while taking a picture.

with the robot.

4.6 MASHIv03: Robot Face

In a social context, a robot should communicate as a human does, *i.e.* through verbal and non-verbal language, mainly to receive and express emotions. In the HRI process, the face of the robot plays an important role [50], and facial expressions are a suitable means to express emotions. Thus, this section presents the development of a robot's face. To do so, a set of animated facial expressions to be shown on the screen of the MASHI robot, a telepresence robot with a wide-screen display mounted on top, were developed.

The analysis and requirements of the MASHIv03 will be detailed in Subsection 4.6.1. Then, the implementation of the robotic platform MASHIv03 is explained in Subsection 4.6.2. Finally, discussion and conclusion about the robot implementation is discussed.

4.6.1 Analysis and Requirements

The use of facial expressions in a robot is very effective for human-robot social interaction, since it can allow the transmission of emotional states, improving the engagement of humans in the communication with the robot [51]. In addition, the fact of having an animated face instead of the face of the operator, allows to carry out studies of HRI using the Wizard-of-Oz technique. Another point to analyze is how realistic the animated face should be. An study from [52] shown that our brain developed in a way in which we now react to emoticons in the same way as we would react to emotion expressing real faces. This means that even very abstract ways of expressing emotions can be understood and are processed in a natural way by humans. Therefore, the animated face to be implemented should have only the necessary features to express the emotional states of the robot.

Also in this section, based on the previous HRI experiences, several improvements will be explained.

4.6.2 Implementation

Mechanical Structure and Appearance

From the HRI experiences carried out previously, some shortcomings were detected in the mechanical design and appearance of the robot head.

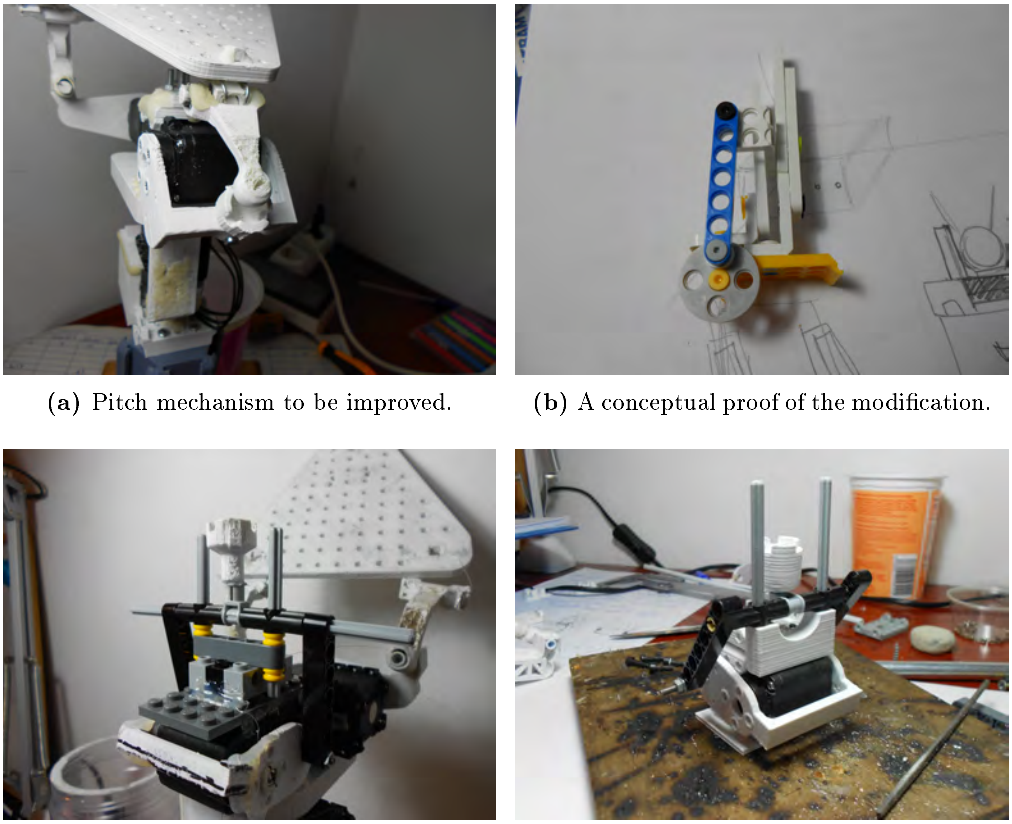
One drawback was the mechanism of vertical movement of the head. In this case, the servomotor shaft loaded with the full weight of the head (see Figure 4.15a), causing it to reheat and block later. We thought of a mechanism in which the head could rest, that is to be in a rest position, without the motor being forced. A conceptual proof of this mechanism can be seen in Figure 4.15b. Based on this conceptual design, we proceeded to adapt this mechanism using pieces for building models, as shown in Figure 4.15c. In the end some component connection pieces were designed in 3D printed while others were reused (see Figure 4.15d).

Another mechanism that presented problems was the articulation of the head, since there was a lot of friction between its parts (see Figure 4.16a). The joints of the head were redesigned to the type of universal joints, as shown in Figure 4.16b.

Some improvements in the appearance of the robot head have also been implemented. For example, the fastening elements to the head structure were located on the front (see Figure 4.17a). We think that in the front of the head (face area) is where people are most fixed at the time of interacting, we must have special emphasis on their appearance. Therefore, in the redesign of the face, the fixation elements have been placed on the sides of the head.

In order to the screen can be seen more as an integrated element in the head of the robot, and that can be perceived as the face of the robot and not as a tablet (shown in Figure 4.17b), a stylized frame (shown in Figure 4.17c), to give emphasis to the eyebrows and the outline of the face of the robot in general, was designed.

The junctions of the case seen in the sagittal plane (see Figure 4.17a) have been eliminated with a new organization of the parts that make up the housing (see conceptual model in Figure 4.17d). The head in general have a more organic and modular design, allowing the rear part of the head to be disassembled to gain access to the head mechanics if required. In Figure 4.17 you can see the final design of the appearance of the head of MASHIv03.



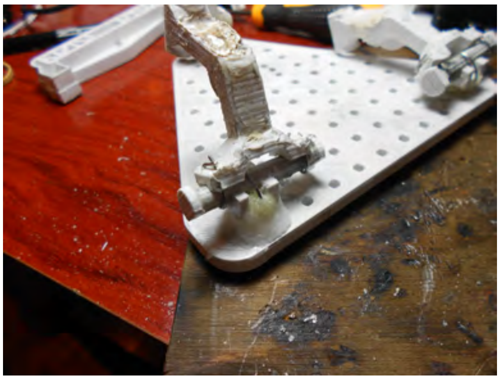
(a) Pitch mechanism to be improved. (b) A conceptual proof of the modification. (c) Modifications introduced. (d) Final disposition.

Figure 4.15: MASHI v03 head mechanical improvements.

Sensors, Actuators and Controllers

A web camera coupled with an omnidirectional lens was added to the top of the head (Figure 4.17e). The objective of this camera is to have a more complete view (360 degrees) of the environment surrounding the robot (see Figure 4.18).

In the movement control of the head a data frame format was used, composed of the operation command, the position of the servo pitch, the position of the servo yaw and the position of the servomotor roll. Thus, several functionalities were created for the movement of the head, such as: rest, zero, position and wake. The rest function deactivates the servomotors, leaving them at rest. The zero function consists of locating the position of the head of the robot to an initial position, which would be seen the head of the robot facing forward. Function wake allows reactivating the head servomotors.



(a) Forced joints.

(b) Redesigned parts.

Figure 4.16: MASHI v03 head joint improvements.

The position function allows us to fix the position of the head given the values of pitch, yaw and roll.

New improvements were also made to the control board of the base movement. An algorithm was implemented to control the data frame of the base movement, which contains an operation code, left wheel speed and right wheel speed. With this the speed of the movement of the base can be controlled. Likewise, the emergency stop functionality was added, which is a special operation code in the data frame that causes the engines to stop immediately without any type of PID control. Another functionality is the rearmament of the motors, which consists in the setting of variables so that the speed control of the motors comes back into operation.

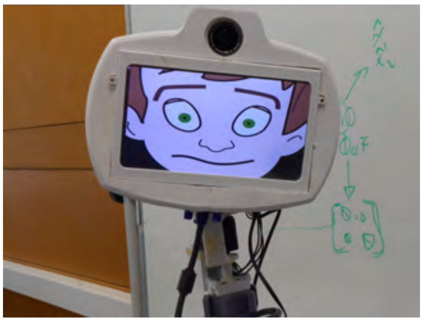
Information and Communication Technologies

Communication protocols were incorporated to control the movements of the robot's head and base, which were explained in the previous section. The code was developed to work with a second camera on the side of the robot, the omnidirectional camera (see Figure 4.18).

The implementation of the facial expressions in the robot, called *RobotIcon*, affects both the operator and the robot components, since the operator needs to control MASHI's emotions, which then in turn need to be displayed on the robot.

Since the existing code base knew only the telepresence mode of operation, where the video stream of the operator is transmitted to and displayed on MASHI, both sides

(a) Head's case.



(b) Installed head's case.

(c) A more stylized frame was redesigned.

(d) Conceptual case.

(e) Final version of the robot's head case.

Figure 4.17: MASHI v03 head mechanism improvements.

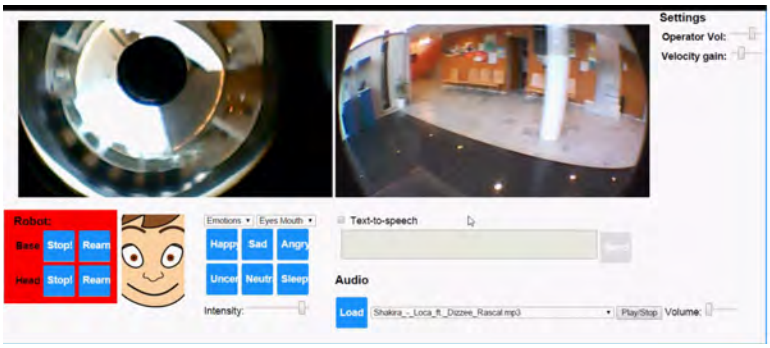


Figure 4.18: MASHI v03 operator's GUI.

needed to be adapted to now support a second, new mode of operation, the telerobot mode, where an artificial face is being displayed on MASHI and it's appearance, in terms of the emotions being displayed, can be controlled from the operator side.

Two main tasks are performed at the operator side, controlling what is displayed on the MASHI's face and comunicating it the robot side. Already at the current stage of the project there are a lot of choices for the user and it was a challenging task to design a graphical user interface which would not be overwhelming for the user. The GUI is built using HTML, JavaScript and CSS.

The existing code was extended to support the control of the emotions as well as toggling between the two modes of operation, *i.e.* the telepresence and the telerobot modes.

To implement the telerobot mode, and allowing for switching between the two modes without a user intervention required on MASHI, this architecture needed to be changed. Therefore, a toggling mechanism was implemented within the MASHI view. This mechanism is encapsulated in the *fullscreen.js*-file which is being loaded by the HTML page and is available through a global variable *fullscreenControl* on the robot side. The initial view remains to be the known menu view, but by calling `setVideoFullscreen` function, the menu view is hidden and the operator's video stream fills the whole page. By calling `Fullscreen()` function, the new RobotIcon view can then be activated, hiding the video completely and displaying only the artificial face on the full available screen.

Therefore, if MASHI now is started, the operator only set the web browser on the robot's screen to fullscreen (with MASHI's current setup e.g. by pressing the `F11` key). After this, all toggling of the operation modes can now be coded from the operator's

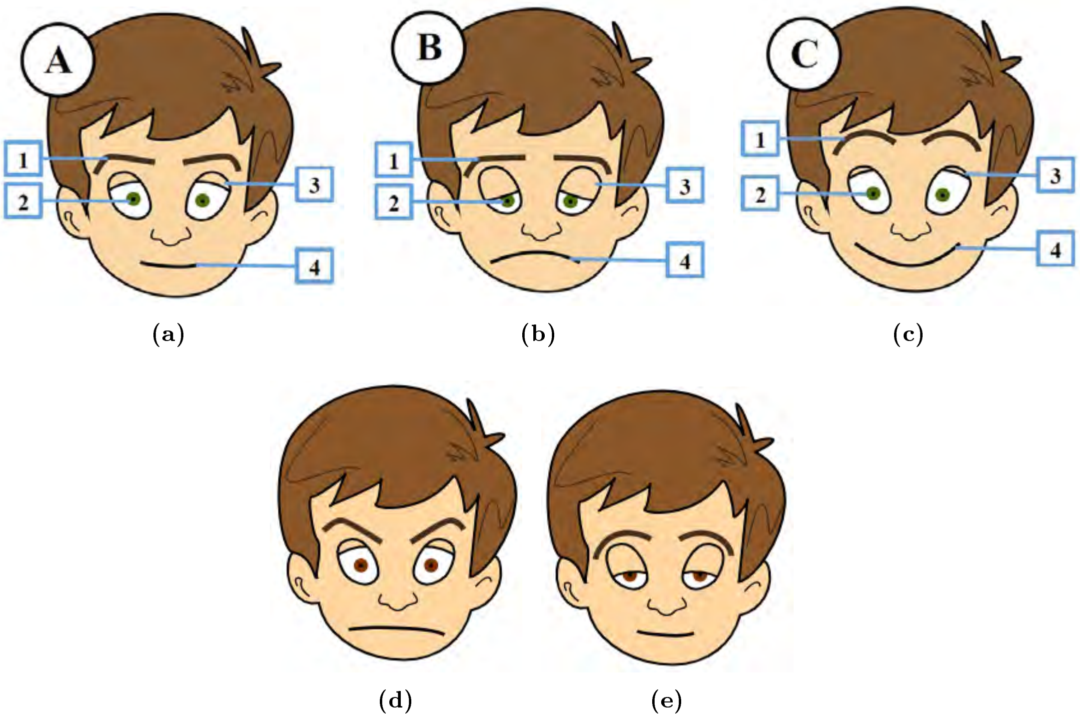


Figure 4.19: Different expressions of the artificial face for MASHI with indication of the animatable regions. (a) shows a neutral expression, (b) a sad emotion, (c) a happy emotion, (c) a angry emotion and (c) shows a sleepy face. Within the face, eleven different attributes can be altered for changing the emotion: (1) the shape and angle of each eyebrow, (2) the direction and intensity of the gaze of each eye, (3) the height of each eyelid, and (4) the shape of the mouth.

side.

As explained in Section 4.6.1, the artificial face on MASHI should help users to relate to the robot, to feel empathy and strengthen social bonding. Therefore, it is crucial that the face looks and behaves human alike. Yet, we deliberately decided to use a face which is *alike* a human face, but in fact is not a real human face.

We decided against actual pictures of faces and for simplified, comic-styled face. This not only allows for easier recognition of the expressed emotions, but also avoids conveying a wrong message – that the robot would have abilities only human beings have. Figure 4.19 shows the face we developed for MASHI. It has eleven attributes which can be altered to change the displayed emotion:

- the shape and angle of each eyebrow,
- the height of each eyelid,
- the direction and intensity of the gaze of each eye, and
- the shape of the mouth.

For being able to display such graphics within the web browser the MASHI interface runs on, different technologies can be used. The main three technologies we took into consideration were:

1. Interchangeable static images,
2. The HTML5 canvas[1] element, or
3. Using a scalable vector graphic (SVG)[2]

Since we early on decided that for a more realistic emotional representation, the transitions between the emotions are an important aspect, the option of using static images was soon rejected. While allowing manipulation of the image via JavaScript, the HTML5 canvas element does not provide a scene graph. This has the effect, that the canvas does not know about any elements being represented – everything is just a pixel – which makes the manipulation of certain elements extremely complicated. Therefore, we decided to use SVG as the underlying technology.

SVG images are being defined by XML[3] files. Within an SVG file, paths, basic shapes – such as circles, ellipses, or rectangles –, or text can be defined. Attributes, like fill color or stroke width, to these elements can be defined for each single element and each element within the SVG can be directly addressed and manipulated.

For easier manipulation of the SVG DOM[4] tree, we employed the open source JavaScript library Snap.svg[5]. With this library, the properties of the elements within the graphic can easily be manipulated and animations can automatically be generated between two transition points.

[1]Canvas is an HTML element, which represents a bitmap surface and can be altered using JavaScript. See `http://www.w3.org/TR/html-markup/canvas.html` for more information.

[2]Find more on SVG here: `http://www.w3.org/Graphics/SVG/`

[3]Extensible Markup Language

[4]Document Object Model

[5]Snap.svg homepage: `http://snapsvg.io/`

Choosing SVG as the underlying technology also allows us to easily change other attributes within the image – such as eye – or hair color. Also adding a completely different face, for example a female character, later in the development process is possible without having to dramatically change the existing logic. Such a new face would simply need to be a separate SVG file with the same structure as the existing file.

The `roboticon.js` library was implemented to encapsulate the logic for the face manipulation. It has to be included on the HTML page displaying the RobotIcon. The RobotIcon SVG graphic is loaded into the HTML page using the HTML5 `object` element. When loaded, the `roboticon.js` library searches for the `object` element with the `roboticon` class set. This element then is being manipulated by the library. To now interact with the RobotIcon, the library exposes the globally accessible variable *RobotIcon*, which in turn exposes methods to animate the RobotIcon.

We have decided to use JSON[1] as the description language for the facial expressions. By calling `parseAndApplyJson` function new facial expression values are applied to the displayed RobotIcon and generates a smooth animation between the previous and the new emotion.

To make the RobotIcon seem more human alike, we have implemented a time interval based eye blinking. Blinking actions improve the sensation of being observed and can be used to provide a more effective impression of social behavior [53]. This interval can later also be used to even better express emotions, like blinking more often when a nervous emotion is shown.

Based on the described architecture in Section 4.6.1, we had to design a way to communicate the new commands from the operator side to the robot. Building on the already existing code, this is being done with the NodeJS socket connection. To now transmit emotions, we have, as already mentioned in the previous section, decided to use JSON as a way of describing the parameters of the facial expression. Since JSON is represented as a string it can be transmitted using the already existing architecture.

In the graphical user interface (GUI) on the operator side, a small window on the top left corner shows what actually is being displayed on MASHI. This way, the operator can assess what users actually see when interacting with the robot. Beneath that, the operator can choose between different operating options. In the first option, *Menu* option, MASHI shows a menu where the user can select different settings or send and

[1] JavaScript Object Notation, find more under: `http://www.json.org/`

receive text messages to or from the operator. Another option, the *Operator* option, MASHI displays the video, that is recorded of the operator in fullscreen (see Figure 4.18). The third option displays the artificial face on MASHI's screen. In this mode, the operator can choose, if the whole face or only the region around the eyes should be displayed. Showing only the eye region of the face makes the appearance more consistent with the hardware of MASHI, because of MASHI's widescreen display and the speaker being located under the display. A discussion of which mode should rather be used will be presented in section 4.7.

Despite the display mode, the user can choose from six different emotions (happy, sad, angry, uncertain, neutral, sleepy) to be displayed and adjust the intensity of the current emotion. The neutral emotion can be used as the basic emotion during communication. The happy, sad, angry and uncertain emotions are very expressive and useful to show reactions to specific inputs. Sleepiness corresponds to a mode where the operator is not commanding the robot, and therefore, the robot can't interact. The intensity slider gives the operator an additional degree of freedom.
The operator can also control the robot's movements with either four different buttons or alternatively four keys on the keyboard. The keyboard should be preferred, when the operator is using a computer. The buttons can be used, when a smartphone or a tablet is used to control the robot. The robot can only be controlled, when the *Control robot* mode is turned on. If this mode is turned off, it is possible to send text messages to MASHI's display. The largest window of the GUI shows the video record from MASHI's webcam. It has to be large, because it makes navigation easier.

Communicative Skills

A substantial improvement in the robot's nonverbal communication abilities skills is the incorporation of the robot's facial expressions. This allows to transmit emotional states of the robot depending on the context and the situation of the interaction. The operator can send two types of visual information: A video stream showing his face or an artificial face, that is controlled by the operator.

If video stream is the operator's face, then any type of non-verbal communication can be shown. The artificial face can express several basic emotions mainly by using the mouth, eyes and eyebrows. The emotions evaluated here are: happy, sad, angry, neutral, and sleepy. While the first three are arguably emotions which may effectively be used within a conversation and neutral is the facial expression made when no emotion is being displayed, the sleepy state was introduced to be able to display that MASHI's operator may be absent at the moment. This way, users may more easily understand why the robot does not respond immediately.

4.6.3 Discussion and Conclusions

In the development of MASHIv03 several improvements were implemented both in the mechanical part and in the programming part.

From the mechanical point of view, the mechanisms of movement and articulation of the head were improved. It also improved the appearance of the head in general, giving a more stylized appearance.

For better control of both the movement of the base and the head, a dataframe format were composed of operating commands and movement commands either for the wheels (speed commands) or for the servomotors (position commands). It would be necessary to implement some kind of data encryption as a security measure, especially in cases of HRI where the operator and the robot are linked by a public network.

In this version it was also possible to incorporate a webcam with an omnidirectional lens but its use consumed many resources of the robot's computer so it made it slow in its operation.

In the context of this work, an artificial comic face, called RobotIcon, have been developed. One goal was to give MASHI a face of its own, with which the robot can express emotions. These emotions can be controlled by a human operator through a browser interface or they can be produced by MASHI autonomously. It is possible to display six different emotions and adjust the corresponding intensity using the face. As future work is the generation of emotions by MASHI automatically through sensors that detect the state of mind of the operator (*e.g.* through computer vision)

The architecture of the system has been designed very generic, which gives much space for several future improvements. One possible improvement would be to add more emotions to allow a more distinct facial communication, which could be realized easily by using our interface for generating emotions. Another improvement could be customize the appearance of the robot's face.

Despite that, alternative ways to command the robot should be evaluated. One possibility would be to generate emotions without buttons, but by representing the operator's emotions autonomously.

In the current stage the robot is still very dependent of the input from the operator. In many situations, it could be possible to make the robot react autonomously (e.g. automatic user emotion detection), which would relieve the work of the operator. What is expected is a robotic platform with different possibilities of representation, with a design that evolves from a telepresence robot to a telerobot with its own internal states.

Using the eyes to focus the user or other objects could also make the robot more socially present. Also, gestures could be employed to emphasize emotions.

In the following section an evaluation of the robot face was detailed.

4.7 Evaluation of MASHI v03

In this evaluation we focus on to which extend the emotions that were created are recognizable. In order to reach wide public and scale the study, the use of an online questionnaire has been decided. The study also tries to answer the question what difference the displaying of the full face opposed to the eye-only region makes in terms of recognizability. This is an important information which could contribute to whether to show the full face with the mouth or whether to show only area around the eyes and therefore whether to suggest an modification in the appearance of the MASHI's face.

4.7.1 Data Analysis

The study was conducted online through a questionnaire on Google Forms[1]. The survey was announced by social networks. It was convenient for the participants and it gave us more information than we would have been able to gather in a manually conducted study. Apart from the basic information about participants, the questionnaire had four major parts. Each participant was asked for gender, nationality, age group and whether the participant have prior knowledge about Human Robot Interaction or have worked in the field of Human Robot Interaction. In each of the four major parts were participants presented with pictures showing five basic emotions and were asked to recognize them.

[1]Google Forms: `http://www.google.com/forms/about/`

The basic emotion shown on the pictures were namely happy, sad, sleepy, neutral and angry.

The main study then had four parts.

In the first part, participants were shown pictures displaying only the area around the eyes while the mouth was not visible (see Figure 4.20a). Participants were asked to recognize emotion shown on the picture and write it into a free-text field.

In the second part, participants were shown exactly the same pictures as in the first part but this time the participants were presented also with a list of our five basic emotions. Participants were asked to select how strongly the picture expresses each emotion on scale from 1 to 7. The full scale from 1 to 7 was translated into “not at all”, “very little”, “a bit”, “some”, “visibly”, “strongly” and “very strongly” respectively (see Figure 4.21).

The third and fourth part were analogous to the first and second, with the difference that now the full face including the mouth was shown (see Figure 4.20b). This way, we can see how important the mouth is for recognizing the correct emotion. The drawback of this ordering is that participants were primed in the free-text part three to write the emotions which they saw as the choices in part two. This is a trade-off we chose deliberately, because if we would have chosen an ordering in which first the two parts are free text followed by two parts with choices, we would prime participants by showing them mouth before letting them choose how strongly are emotions expressed only by eyes.

4.7.2 Results

During 6 days, 102 participants took part in the study. 33% of the participants were female and 67% were male. 22% of our participants had a prior knowledge in the field of human robot interaction. Our study attracted people from 16 countries but the most participants come from Germany, Slovakia and Spain which together counts for over 80% of our participants. The distribution of nationalities is shown as pie chart in figure 4.22a. The most numerous age group of our participants was from 21 years to 25 years. The detailed distribution of age groups of our participants is shown in figure 4.22b.

From a first analysis of the free-text answers a three to four times higher diversity was given when participants were presented with the eyes only part in comparison to the full face part. Table 4.4 shows the percentages of people who entered a word in the

(a) Eyes only **(b)** Full face

Figure 4.20: Face area shown in the questionnaire.

free-text field recognizing the displayed emotion correctly, with eyes only part in the first line and the full face in the second line. At a first glance, very high values for the full face, and slightly smaller values for the eyes only part can be observed. An outlier can be found when looking at the happy emotion with the eyes only being displayed.

Table 4.4: The percentages of people who correctly recognized each emotion in free text parts.

	Happy	Sad	Sleepy	Neutral	Angry
Eyes only	1%	51%	63%	28%	65%
Full face	92%	97%	86%	86%	97%

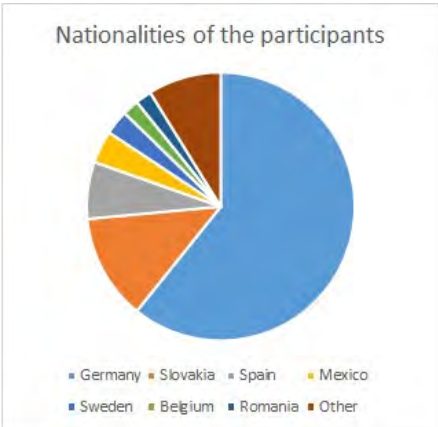
Figures 4.23a and 4.23b, as well as 4.23c, 4.23d, and 4.23e show the average values of how strongly each emotion was said to be present in each picture. The higher the bar is, the more the emotion is being recognized within the picture. The blue columns show the results from the eyes only part while the orange columns correspond to the full face expression.

For the happy emotion there was the biggest difference between showing the mouth or not. Actually over 80% of participants wrote surprise instead of happy in free text

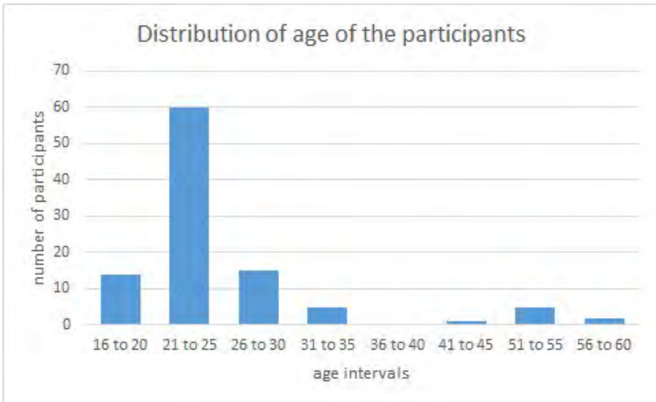

18.1 - Please select how strongly the image expresses the emotions below on a scale from 1 (not at all) to 7 (very strongly) *

	1 - not at all	2 - very little	3 - a bit	4 - some	5 - visibly	6 - strongly	7 - very strongly
Happy	○	○	○	○	○	○	○
Sad	○	○	○	○	○	○	○
Sleepy	○	○	○	○	○	○	○
Neutral	○	○	○	○	○	○	○
Angry	○	○	○	○	○	○	○

Figure 4.21: Choices of emotions and intensities in the questionnaire for part two and four.



(a) Nationality.

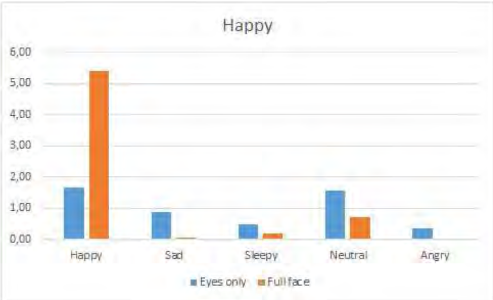


(b) Age.

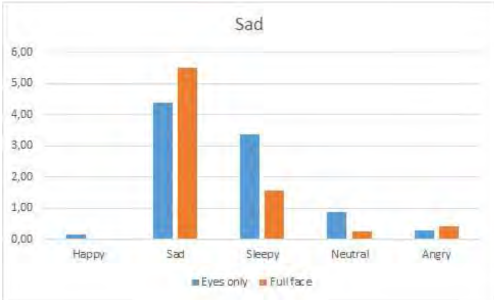
Figure 4.22: Results group distribution: (a) by nationality and (b) by age.

answers for eyes only part. By showing the smiling face the performance improved rapidly from 1% to 92% and only 3% of participants wrote surprise when there was a mouth. One can see that it is difficult recognize happy emotion also by looking at the big difference in size of the bars in figure 4.23a.

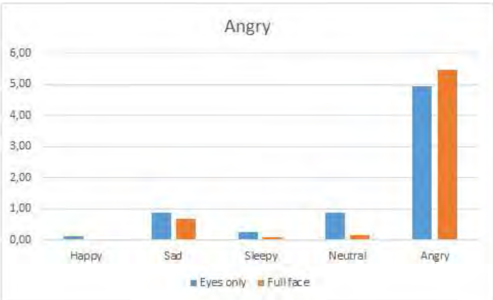
The sad emotion turn out to be easily confusable with sleepy emotion. Although the sadness was the most popular answer in free text part, it was closely followed by the tired and bored. Again the addition of the mouth to the sad eyes removed the doubts of which emotion is expressed. We can see it both from the improvement by 46% in free text part and also by much smaller orange bar compare to blue bar for sleepy emotion



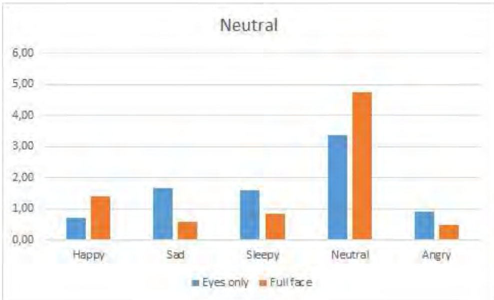
(a) Happy



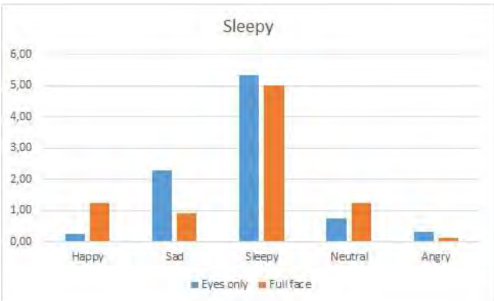
(b) Sad



(c) Angry



(d) Neutral



(e) Sleepy

Figure 4.23: How strong is each emotion expressed.

in Figure 4.23b. Nonetheless we can see that the sad emotion is strongly present in both cases.

For the sleepy emotion there was the smallest difference between the eyes only and the full face. It is the only one emotion which had stronger presence of sleepiness without the mouth than with mouth as can be seen in Figure 4.23e. Possibly this result is that the biggest gesture of being sleepy is to be with the eyes closed or partially closed, rather than the gesture of the mouth.

The neutral emotion earned the position of the worst recognizable emotion. In the free text part for eyes only there was the highest variety of answers and all together over 60 different ones. When the participants were presented with the choices (see Figure 4.23d) the performance was much better as others emotion were not present there either.

The angry emotion was on the other hand the best recognizable emotion. Almost in all metrics it has the highest performance. Only from eyes region 65% of participants correctly recognized the emotion and with the addition of the mouth the result was close to hundred. Also from Figure 4.23c can be seen that the presence of other emotion is negligible compare to the angry emotion.

To sum up we can mention that in general the robot emotions can be recognized from the whole face. The addition of the mouth was crucial in order to recognize the happy emotion.

4.7.3 Discussion and Conclusion

A survey was conducted to evaluate the recognition value of the facial expressions as well as to evaluate whether displaying the full face or rather just the eye-region affects this recognition rate. The aim of this work is to analyze the degree to which people can recognize comic-like facial expressions of a robot, especially when these facial expressions include or not the mouth.

Two different types of views were compared, where the entire face or only the eye region of the face was shown. The evaluations has shown, that it is highly beneficial to display the whole face in order to transmit all emotions properly, especially for showing happiness. However with very careful design and further evaluation with users it is possible to express at least some emotions using only the eyes.

As a future work the comparison could be made but with the sleepy emotion of the whole face modifying the type of mouth (some kind of open or yawning mouth), instead of just a single line which is currently used as mouth.

In order to have an external validity, and therefore avoid bias between showing full face or only eyes, we could have randomly assigned half the participants first eyes and then full face, and the other half, full face and then eyes, both with free choice. And then the same, but with the 7-point scale. This way you ensure you have equal number of participants exposed to the different ordering which allows you to evaluate as well the influence of either seeing full-face before eyes-only, or the other way around, and also possible combinations of free-text vs. scale.

The testing of both the implementation of the MASHI v03 robotic platform as well as the evaluation of facial expressions have been described.

Then, the last phase of the platform presented to date is presented. From the previous experiences it can be mentioned that the provision of arms to the robot will improve the HRI in different aspects.

4.8 MASHI v04: Robot Arms

In this section the implementation of the MASHIv04 is presented. In Section 4.8.1 the requirements for MASHIv04 are presented. The implementation of the different parts of the platform are explained in Section 4.8.2. The discussion and conclusions of the results obtained are finally presented in Section 4.8.

4.8.1 Analysis and Requirements

As discussed in previous HRI experiences, the arms of a robot can help in the generation of non-verbal language or in the generation of interactive behaviors. Thus, the arms of a robot could generate attentional gestures [34], where through the movement of the elbows, the robot could point to some place, object or person. In the case of a robot guide, *e.g.* the arms of the robot can serve to indicate the different works of an exhibition. In the case of interactive behaviors, there are some behaviors that can be recreated with the help of arms, such as being accompanied by the hand (Figure 4.11a), holding a hand(Figure 4.11d), hugging (Figure 4.14c), among others. In our case, the arms of the robot aims to improve the channel of non-verbal communication and the production of certain interactive behaviors, but not to perform any kind of physical manipulation of the environment.

Therefore, the main objective in this version is the design and construction of articulated arms for the robot considering its anthropomorphic proportions. The number of joints proposed will depend on the service provided by the robotic platform.

In a first sprint, having the robot as an exhibition guide, articulations for the elbows were motorized, in order to produce attentional gestures in HRI (see Section 5.3).

In a second phase, a motorized articulation for the left shoulder was required, in order to rise its arm and give the robot the functionality of taking pictures with the people (selfies). See Figure 5.4 for an illustration. The main developments in MASHIv04 are reviewed below.

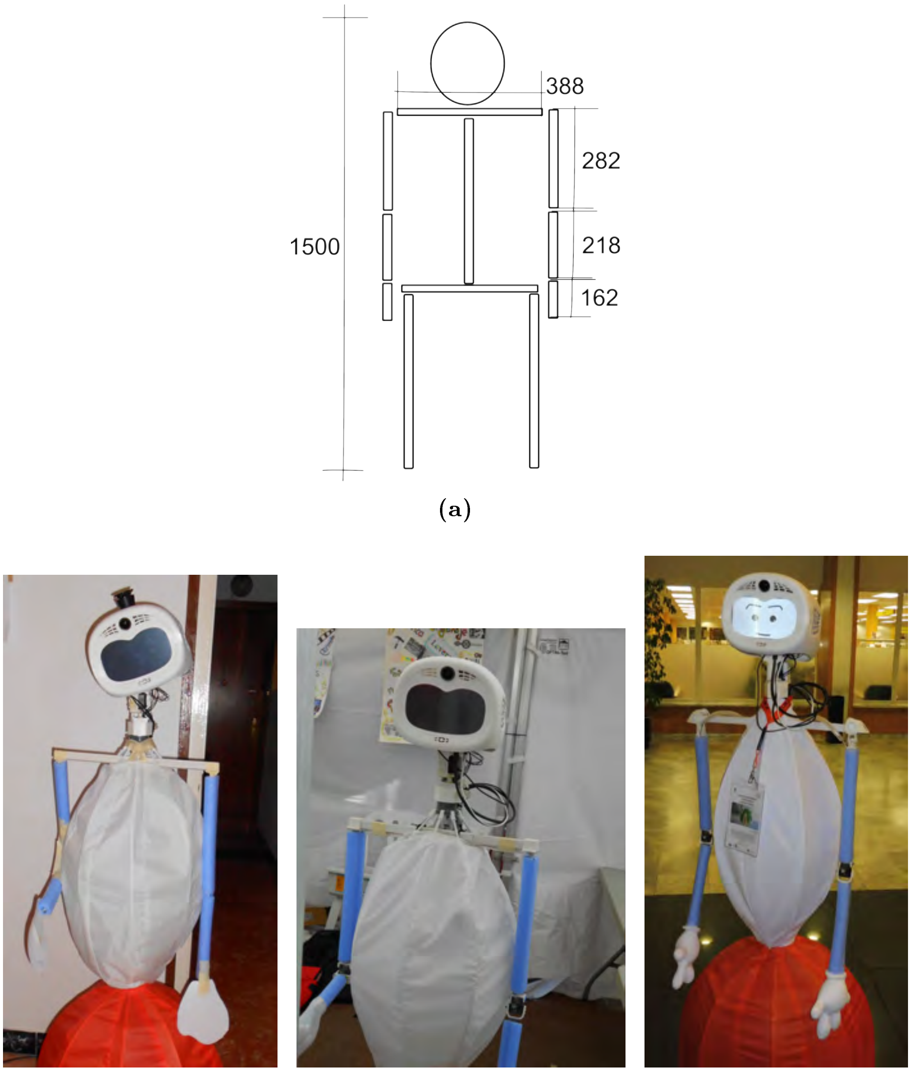
4.8.2 Implementation

Mechanical Structure and Appearance

With the aim of constructing two articulated arms in the robot to improve the interaction with people, initially a conceptual design of the arms was made maintaining the anthropomorphic proportions of the robot according to its height (1500 mm), at it can be seen in Figure 4.24a. For this purpose, a cable duct was used as part of the clavicle, PVC rods to emulate the bones, polyethylene protection profiles for skin and cardboard to represent the hands, as shown in the Figure 4.24b.

In a first version, for each arm two PVC rods were used, one for the arm and one for the forearm, a motorized articulation for the elbow and fixation connectors printed in 3D (see Figure 4.24c). A variant was to reinforce the arm and forearm with two rods in each one, as well as adding a universal joint at shoulder level, as shown in Figure 4.24d

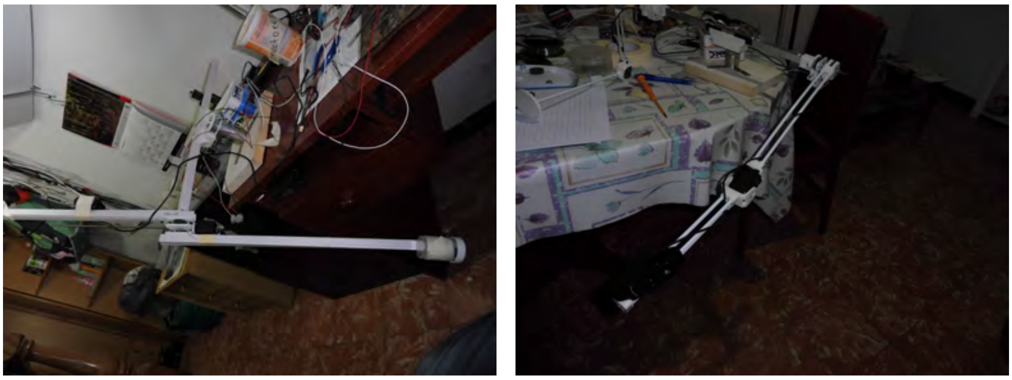
In a second sprint the left shoulder of the robot was motorized and a web camera was adapted in its hand so that the robot could raise its arm and take photos with the visitors (selfies). Details of the experience can be found later in Section 5.4. To do this, the alternatives were analyzed and conceptual tests were carried out for the shoulder mechanism. One option was to put an engine with greater torque and that the motor shaft coincides with the shoulder axis; however the high price and the knowledge that the engine was going to be in continuous operation and with the risk of being forced constantly, there was a high probability that the engine had been damaged. Then two



(a)

(b) (c) (d)

Figure 4.24: MASHI v04 first sprint: (a) Conceptual design; (b) First approach; (c) Second approach; (d) Third approach.



(a) Gearbox rise the arm as a counterweight.

(b) Gearbox raised directly the arm.

Figure 4.25: MASHI v04 second sprint.

prototypes were tested, both using a worm-type gearbox. In the first, the gearbox allowed a bar to rise as a counterweight to the arm, making it lift (see Figure 4.25a). In the second, the gearbox raised the arm directly (see Figure 4.25b).

Sensors, Actuators and Controllers

For the implementation of the arms, the Dynamixel XL-320 servomotors were used in the first tests, due to the low torque requirement because the low weight of the arms. For the MashiSelfie version, the Dynamixel AX-12A servomotors were used for their best performance in torque.

The same OpenCM controller for the head was used to control the servomotors of the arms. In the programming of the controller variables and algorithms were added for the control of the servomotors of the arms: movements of joints for the elbows and articulation of the left shoulder.

Information and Communication Technologies

In the communication platform of the robot several implementations were made: on the operator's side, the movement controls of the arms and the left shoulder were added by means of the keyboard as well as the transmission of the movement instructions to the server side (robot). On the server side, it was added the reception of the movement instructions of the arms coming from the operator and its transmission to the OpenCM.

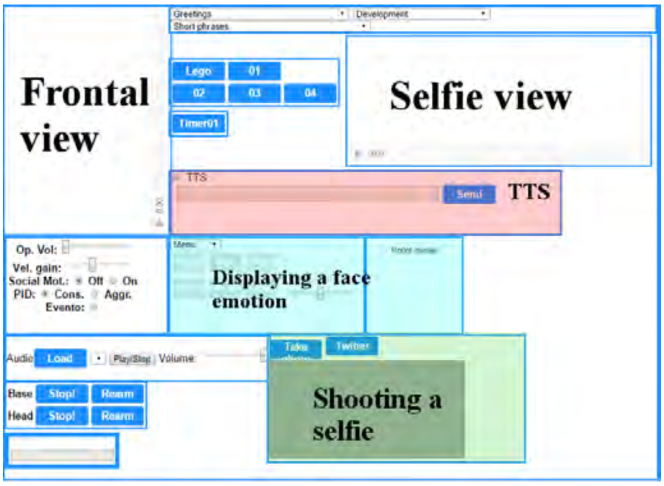


Figure 4.26: MASHI v04 operator's GUI.

In addition, for the experience in Section 5.4 a window was added in the GUI to show the video of the camera placed in the hand of the robot, and another window to show the photo taken from this camera (see Figure 4.26). Another functionality that was implemented was the service of sending the photo taken by the robot to the @MashiRobot account of the social network Twitter.

Communicative Skills

The implemented arms expand MASHI's the nonverbal communication channels as well as the behaviors of the robot. For example, raising the elbow, left or right, allows to indicate a tarjet (social agent or stimulus) that is in the projection from the elbow to the hand (attentional gestures). A sequence of synchronized movements of the elbow could give the appearance of shaking hands. Depending on the context, raising both elbows can give the impression that the robot wants to break through or want to hug. Raising the shoulder would give the impression that the robot wants to take a selfie with the users.

4.8.3 Discussion and Conclusions

In this section the development of MASHI v04 was presented, which consisted mainly in the construction and programming of the robot arms. Wrapped in a process of continuous improvement based on the service provided by the robot, different evolutions were made in the design of the robot arms.

The endowing of arms to the robot widens the capacities of nonverbal communication, thus enhancing their social human-robot interaction. Nevertheless, given more degrees of freedom - DOF to the robot arms, the movement control by the operator becomes more complex. New interfaces to operate the robot arms, not only by the keyboard, could be used, for example through the use of depth cameras trough the detection of the operator gestures. In addition, it is necessary to provide greater autonomy in the movement of the arms, to reduce the workload of the operator, and with greater reason as the movements of the arm become more complex.

Chapter 5

Testing and Improving MASHI from Users Feedback

In this chapter we present different experiences of the MASHI robot in natural environments. In a first part, the experiences of the robot as an exhibition guide carried out in a cultural center are detailed, while in a second part the experiences of the robot with the role of promoter in an international trade fair are presented.

5.1 A First Experience in a Cultural Center

In this section we describe an exploratory study of HRI with our MASHI robot in a Cultural Center. With this new scenario and with our robotic platform, in this study we will use direct observation techniques to analyze the behavior of HR groups; likewise we will conduct a survey to know the judgements of the participants on different characteristics of the robot and the interaction. MASHI is tested as an exhibition guide in "Exposicio de Joieria '(2)' " (*Jewelry Exhibition '(2)'*)(Figure B.1).

This section is organized as follows: In Subsection 5.1.1 the study design is detailed. The methods used to collect data are detailed in Subsection 5.1.2. The results of both the observations and the survey carried out are presented in the Subsection 5.1.3. In Subsection 5.1.4, discussion and lessons learned from the results obtained and the overall experience are presented. Conclusions and future work are finally presented in Subsection 5.1.5.

5.1.1 Study Design

This section discusses general issues related to the design and development of the experience in a Cultural Center. The MASHI robot was deployed in the main hall ($6 \times 8m^2$ aprox.) at La Bòbila Cultural Center located in L'Hospitalet de Llobregat, a town near Barcelona, Spain. The field study was carried out for two weeks, from 14th to 30th April, for about 2 hours per day at the afternoon.

Objective

The main aim in this experience is to observe social human-robot interactions *in-the-wild* with a guiding robot in the context of a Cultural Center. In this first experience our specific objectives are:

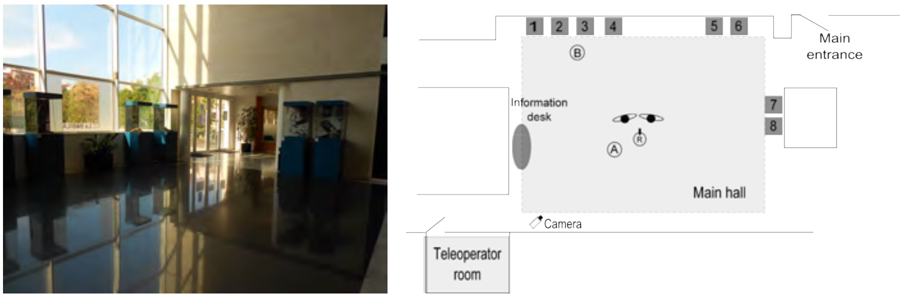
- To study the visitor preferences interacting with the robot
- To study visitor characteristics with the robot in this social scenario
- To describe the group HR behavior in terms of the spatial arrangements
- To have a feedback from people about the HRI experience and about the appearance and functioning of the robot.

Scenario and setup

La Bòbila Cultural Center, in L'Hospitalet de Llobregat, is a three floor building containing multiple facilities for education and leisure: a library, an auditory, different rooms for courses and other activities, and a main hall with temporal exhibitions around it.

The robot was deployed in the main entrance hall at La Bòbila, an area of about 8 meters wide and 6 meters long near the main access from the street and around it there were 8 exhibitors on a jewelry exhibition (see Figure 5.1b and Figure 5.1a); so the exhibition occupies an area where people coming to other areas of the building had to go through anyway.

Two locations (A and B) were defined as the main possible destinations for the robot guide. Point A is the initial location of the robot in the center of the hall and in front of the information desk. Point B represents the exhibition guiding area, comprising eight works of jewelry (see Figure 5.1b).



(a) Main hall at La Bòbila **(b)** Setup

Figure 5.1: Scenario and setup.

The field study was carried out from 14th to 30th April for about 2 hours per day at the evenings matching in time with the exhibition of jewelry called "(2)". No adaptation of the physical environment was implemented to maximize the study ecological validity preserving the natural every-day conditions and routines except from a zenital camera placed in the second floor out of sight of visitors at a height of approximately 3 meters in order to have an overall view of the scene (Figure 5.1b). The operator room was just next the main hall in a private room inaccessible for visitors (see Figure 5.1b).

Procedure

The robotic platform was used in a Wizard-of-Oz setup thus the robot's head movements, displacements, dialogue with visitors and interactions were totally teleoperated by the operator that remains out of visitor's sight inside the operator room.

According to its role the general function of the robot is to enrich visitor's experience by exhibiting itself as an attraction, providing entertainment and guiding people through the exhibition. The robot's role is deployed in three activities: dialogue, entertainment and guidance (as seen in Figure 5.2). From these activities, the first two were open topic, in order to explore the kind of questions asked by visitors or the type of games they propose. The last activity had a predefined text for each of the eight stands. Therefore, it was left to the operator discretion the use of verbal and non-verbal (like the facial expressions and the head's movement) language in order to maintain the interaction around these three activities (except for the script used in the exhibition guide (see Appendix B.1).

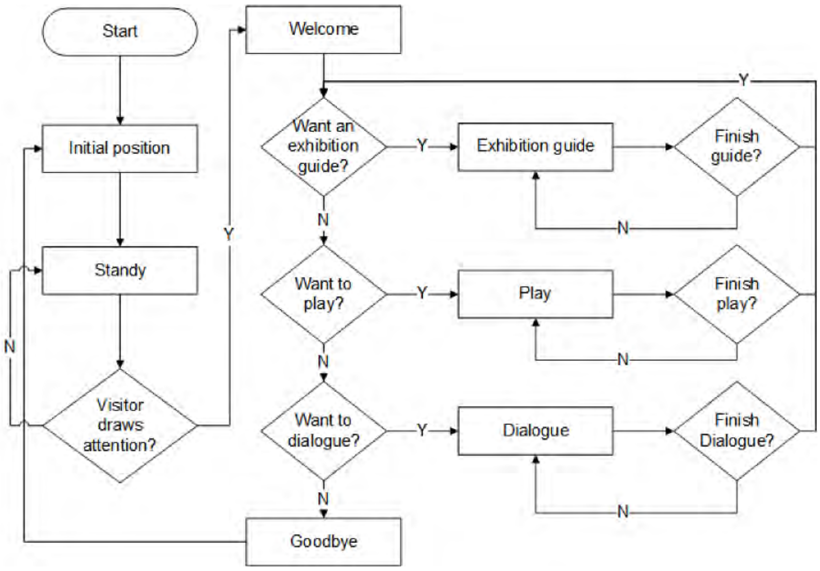


Figure 5.2: Flowchart of robot's role.

In the first instance the robot is in a standby state: *i.e.* robot steady in initial position and facial expression sleepy, as it can be seen in Figure 5.3a. Once one or more visitors come to the robot, it changes to the 'Welcome mode': neutral facial expression and utters an spoken message of welcome (Figure 5.3b). In the case of people wishing the exhibition guide service, the robot will tour and explain each exhibit case (Figure 5.3c-5.3f). Once the tour is completed, the robot ask visitors if they want guidance again or whether they wish to play or have any questions. If any further service is requested, the robot say goodbye and return to its initial position area (Figure 5.3g-5.3h).

No briefing or instruction was given to visitors, and the intervention of technical staff at the local environment was exclusively aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

5.1.2 Data Gathering

In this study we use two types of methods: direct observation and a survey, in order to describe the group HR behavior and the perceptions of the robot and the HRI in general.

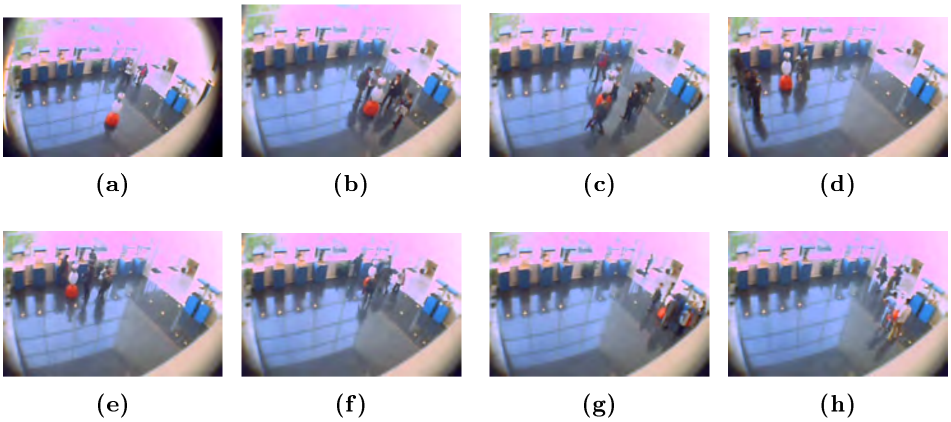


Figure 5.3: A robot guiding during the exhibition (2).

The group description and spatial arrangements categories (see Table 5.1) was performed using human interpretation of verbal and non-verbal communication such as body orientation, gestures and group spatial cohesion. In the context of this study to detect and categorize groups in HRI it has been convenient to use the recordings from the external camera, since it offers a general view of the experimental area. Of these, those video sequences will be selected, called episodes, where some type of the aforementioned interactions can be observed.

Table 5.1: Group characterization and spatial arrangements categories.

Dimensions	Variables	Categories
Group characterization	Size	Single
		Couple
		Triple
		Larger
	Composition	Children (6-11 years old)
		Young (12-17 years old)
		Adult (18++ years old)
		Mixed
Spatial relationships	F-formations	'Via-a-vis' (dyadic)
		'L-shape' (dyadic)
		'Circular form'
		'Horseshoe shape'
		'Side-by-side'
		'Performer-audience shape'
	Proxemic behavior	Intimate
		Personal
		Social

To carry out the survey, a questionnaire was prepared with 22 questions. The first 18 questions are sentences with a Likert scale of 5 options. Then, one question follows with an open answer. The rest of items were activities of drawing and coloring the robot on the next page (see Table 5.2). An example of the questionnaire can be seen in Apendix B.

In general, the questions were focused on knowing the perceptions of the people about the HRI experience. In particular, for the robot as an exhibition guide, about the role played by the robot, about the characteristics of the robot and about the behavior as a group (as shown in Table 5.2).

Since in this case we are based on a series of Likert-type individual questions in which we want to evaluate different aspects of the HRI, we can not use the reliability criterion of Cronbach's alpha.

5.1.3 Results

From observational data

The analysis took into account the recordings of six sessions, with a total time of 481 minutes.

From this data, 32 human-robot interactions or episodes were observed with a total time of 325 minutes, representing an occupancy rate of 67.7% of the total time that the robot was in the hall.

It should be noted that, given the dynamics of the interaction in this public setting, it can be seen different compositions of groups and spatial relationships within the same episode, thus the results shown below reflect an occurrence degree for each category.

A descriptive analysis of the data showed 14 episodes (38.9%) of visitors playing with the robot (10 children, 3 young and 1 adult), 13 (36.1%) episodes of visitors maintaining a dialogue (9 children, 2 young and 2 adults) and 9 episodes (25.0%) of visitors in the exhibition guide (6 children, 1 young and 2 adults)(see Figure 5.4).

Concerning the group composition, the highest percentage of groups that interact with the robot were children (23 occurrences or 69.7%) , 15.2% (5 occurrences) were adults, 12.1% (4 occurrences) were young and 3.0% (1 occurrence) were a mixed group composed by children and young visitors.

Table 5.2: Questions and their focus used in the questionnaire.

	Question	Focused in
Q1	Mashi is a good exhibition guide in La Bòbila.	Robot role
Q2	I like to accompany the robot during the exhibition.	Group behavior
Q3	I like the robot to follow me.	Group behavior
Q4	I like to follow the robot.	Group behavior
Q5	I would like Mashi to be faster.	Robot characteristic
Q6	The movements of the head are adequate.	Robot characteristic
Q7	I like the extra activities that the robot does.	Robot role
Q8	I like the robot to recognize my name.	Robot characteristic
Q9	I like the robot to play music.	Robot role
Q10	I like to dance with the robot.	Robot role
Q11	I like to hug the robot.	Robot characteristic
Q12	I'm not afraid of Mashi, he looks harmless.	Robot characteristic
Q13	I like to get close to the robot.	Group behavior
Q14	The robot looks very heavy.	Robot characteristic
Q15	The robot is very tall.	Robot characteristic
Q16	I like the face of the robot.	Robot characteristic
Q17	I like the facial expressions that the robot does (eg happiness, sadness, etc.)	Robot characteristic
Q18	I would like the robot to have arms.	Robot characteristic
Q19	If Mashi had arms, why do you think they could serve?	Robot characteristic
Q20	Paint Mashi! Color Figure 1 the way you like it!	Robot characteristic
Q21	Would you like Mashi to have arms? Draw some in Figure 1!	Robot characteristic
Q22	Would you like Mashi to have another face, maybe yours? Draw and paint a face to Mashi (Figure 2)	Robot characteristic

(a) Playing (b) Dialogue (c) Exhibition guide

Figure 5.4: Interaction preferences.

Visitors who interact alone with the robot were 9.1% children, 3.0% were young and 3.0% were adult; triples were 3.0% children. Larger groups were composed by 36.4% children, 6.1% young, 12.1% adults and 3.0% mixed ages. Mixed group sizes were formed by 21.2% children and 3.0% young visitors (see tabulated results in Appendix B.2).

Table 5.3: Group composition vs. group size

	Size				
Composition	Single	Triple	Large	Mixed	**Total**
Children	3 (9,1%)	1 (3%)	12 (36,4%)	7 (21,2%)	23 (69,7%)
Young	1 (3%)	0 (0%)	2 (6,1%)	1 (3%)	4 (12,1%)
Adult	1 (3%)	0 (0%)	4 (12,1%)	0 (%)	5 (15,2%)
Mixed	0 (0%)	0 (0%)	1 (3%)	0 (%)	1 (3%)
Total	5 (15,2%)	1 (3%)	19 (57,6%)	8 (24,2%)	33 (100%)

F-formations were encountered during interactions, the dyadic ‘vis-a-vis’ and ‘l-shape’ arrangements observed at 17.6% and 2.0% of the interactions, respectively. ‘Circular form’ was observed at 49.0%, ‘horseshoe shape’ at 13.7%, ‘performer-audience’ distribution at 9.8%, while the 7.8% were ‘side-by-side’ arrangements (see Figure 5.5 and Table 5.4). It is worth mentioning that of the 15 episodes with only circular formations, 6 corresponded to dialogue and 9 corresponded to game and none of them corresponded to the guided tour.

Regarding the proxemic behavior observed, 11 occurrences (26.8%) were in the intimate space, 22 occurrences (53.7%) were in the personal space while 8 occurrences

Table 5.4: F-formations

Item	Occurrences	Percentage
Vis-a-vis	9	17,6%
L-shape	1	2,0%
Circular form	25	49,0%
Side-by-side	4	7,8%
Horseshoe shape	7	13,7%
Leader	5	9,8%

(19.5%) were in the social space. Examples can be seen in Figure 5.5. It should be noted that the intimate distance corresponds in general to children, with a group size of three or more users and with preferences in playing and dialoguing.

From the survey

A total of 15 children took the survey, 4 of which were male, 10 were female and one did not respond. The average age was 10.18 years old and a standard deviation of 1.75.

Robot role As we can see Figure 5.6, the results of the questions related to the role of the robot reveal the following percentages of people who are in total agreement: 80% think that MASHI is a good exhibition guide in La Bòbila (Figure 5.6a), 67% likes the extra activities that the robot does (Figure 5.6b), 64% likes the robot to play music (Figure 5.6c) and %40 likes to dance with the robot (Figure 5.6d).

Group behavior As we can see in Figure 5.7, the results of the questions related to the group behavior reveal the following percentages of people who are in total agreement: 67% likes to get close to the robot (Figure 5.7d), 62% likes to accompany the robot during the exhibition (Figure 5.7a), 53% likes the robot to follow me (Figure 5.7b) and 47% likes to follow the robot (Figure 5.7c).

Robot characteristics As we can see in Figure 5.8, the results of the questions related to the robot characteristics reveal the following percentages of the participants who are in total agreement: 80% like the robot to recognize his name (Figure 5.8a), 80% like the facial expressions that the robot does (eg happiness, sadness, etc.) (Figure 5.8b),

Figure 5.5: Spatial arrangements: (a) 'vis-a-vis', (b) 'circular form', (c) 'horseshoe shape'; and proxemic behaviors: (d) intimate, (e) personal, and (f) social distances.

73% were not afraid of MASHI, they see it harmless (Figure 5.8c), 67% would like the robot to have arms (Figure 5.8d), 60% think that the movements of the head are adequate (Figure 5.8e), 57% like the face of the robot (Figure 5.8f), 43% would like Mashi to be faster (Figure 5.8g), 40% likes to hug the robot (Figure 5.8h), 40% think the robot is very tall (Figure 5.8i) and 20% think the robot looks very heavy (Figure 5.8j)

As we have seen, the visitors positively valued (Q18 in Figure 5.8d) the fact that the robot had arms. On this same aspect regarding the question of if MASHI had arms, why do you think they could serve? (Q19 in Table 5.2) the answers were, among others: *to point out the pictures that explain and to be able to embrace*, *to give hugs*, *to hug*, *to embrace people and dance in front of everyone*, *to dance*, *to give your hand*, *to catch things*.

At the end of the survey, when visitors were asked to paint the robot and draw some arms, there were very creative proposals, as an example you can see Figure 5.9. In the Appendix B.3 you can see all the proposals made by the participants.

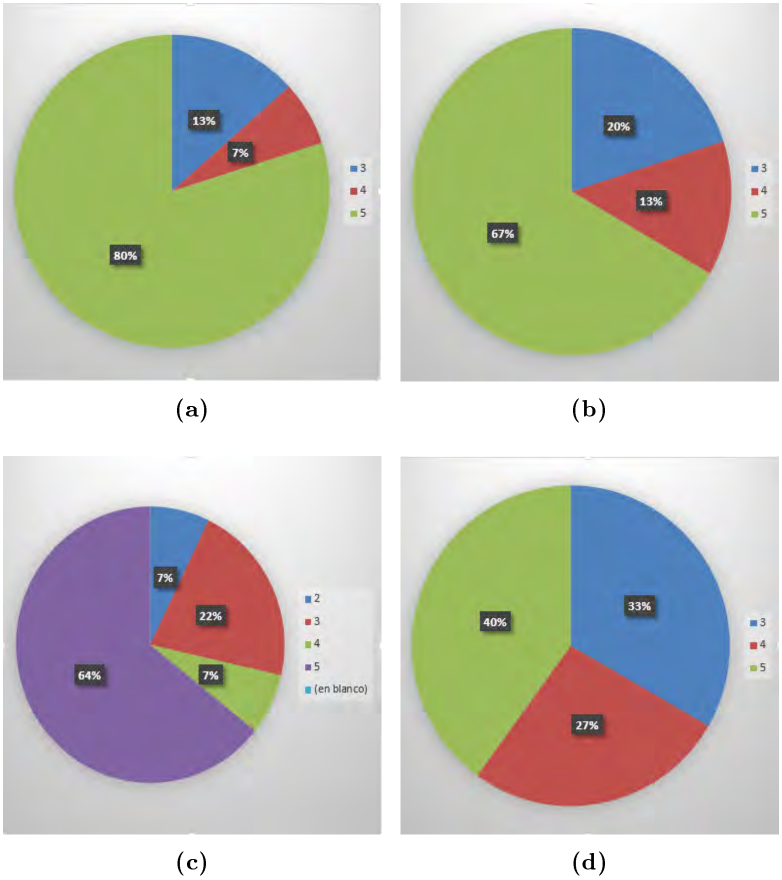


Figure 5.6: Survey results of robot role related questions: (a) Q1 Good exhibition guide?; (b) Q7 Like extra activities?; (c) Q9 Like plays music?; (d) Q10 Like dance?.

5.1.4 Discussion

In the exhibition guide it was considered that once the tour started, it had to be completed before the end of the interaction. It was not taken into account that there could be people who come and go. It should be taken into account for future HRI what the robot should do in those cases.

Leaving open the use of both verbal and non-verbal language (facial expressions and head movement), to a lesser extent in the guided part where there was a script on the presentation of the exhibition, can have an impact on how interactions are performed , affecting the results. This first experience sets the pattern for making such improvements in future HRIs.

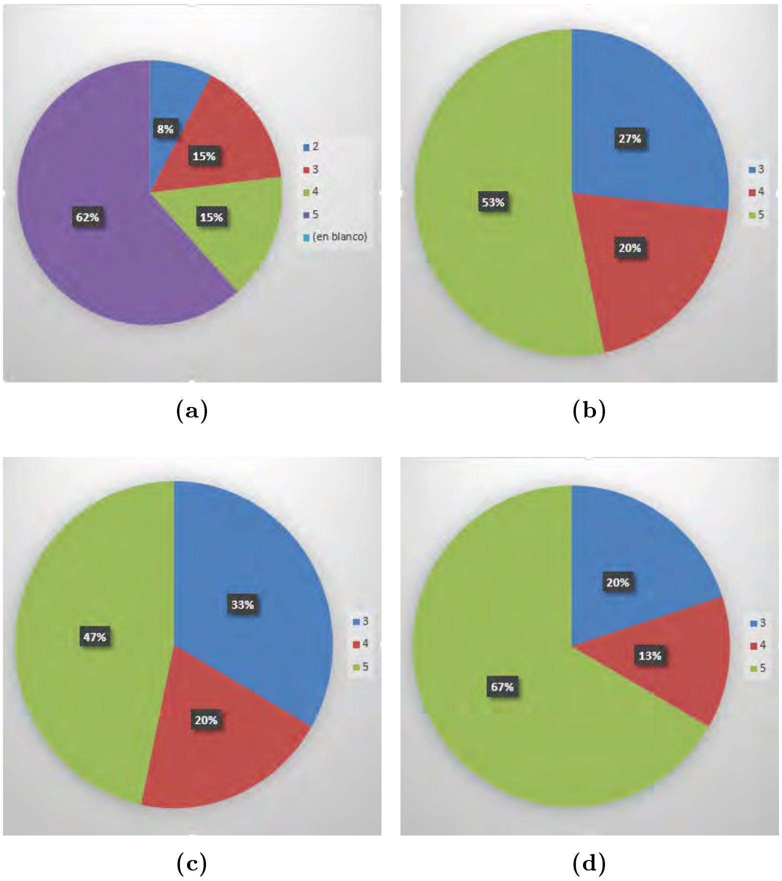


Figure 5.7: Survey results of group behavior related questions: (a) Q2 Like to accompany the robot during the tour?; (b) Q3 Like robot follow me?; (c) Q4 Like to follow the robot?; (d) Q13 Like to get close the robot?.

Due to the highly dynamic nature of this open environment, the groups formed during interactions continuously change both in structure and behavior. The changes observed in groups during interactions were given mainly in their size, their spatial arrangements and their proxemic behavior. It was observed, however, that during interactions the group age don't vary substantially. For example, if a group of children initiated the interaction, although it could vary their dimension and spatial behavior, usually the "group age" was maintained until the end of the interaction.

Unlike the results obtained in [54], where several arrangements were observed during displacements, in this study were few occasions when some kind of spatial arrangement was detected in the exhibition guide. Two factors that can influence this issue could be

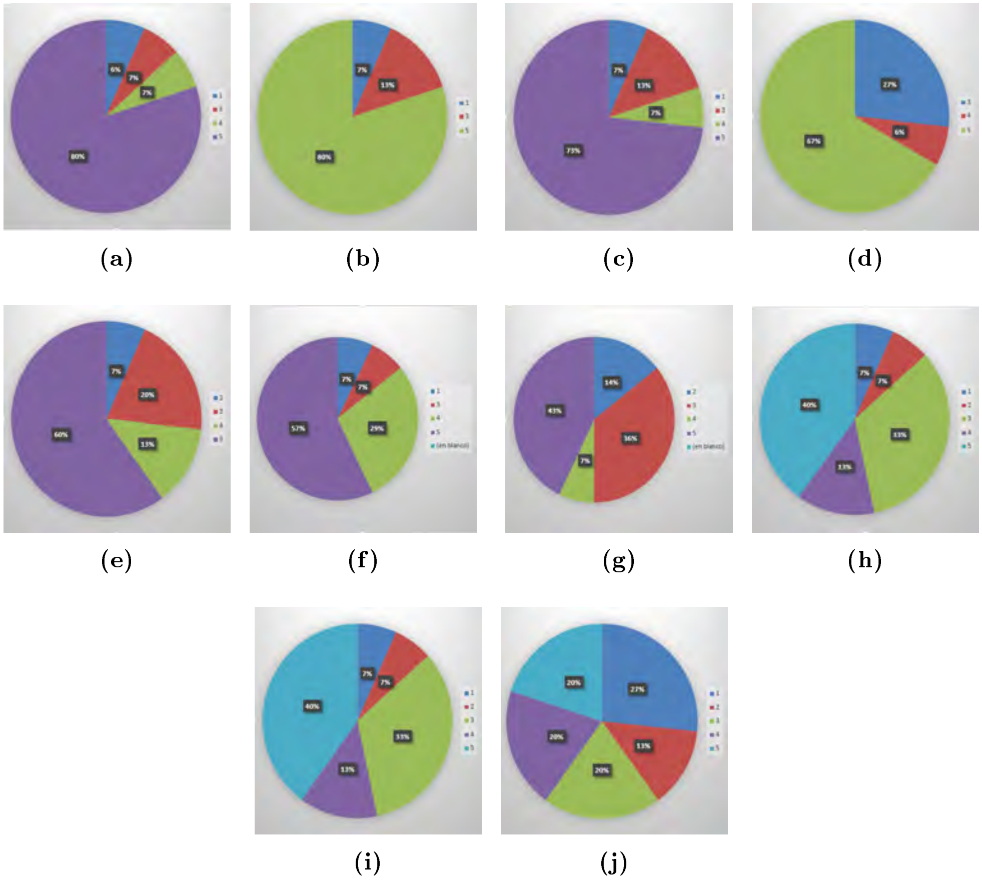


Figure 5.8: Survey results of robot characteristics related questions: (a) Q8 Recognize my name; (b) Q17 Facial expressions; (c) Q12 Feel not afraid; (d) Q18 Had arms; (e) Q6 Proper head movement; (f) Q16 Like face; (g) Q5 That was faster; (h) Q11 Hug the robot; (i) Q15 Is tall; (c) Q14 Looks happy.

physical constraints of the environment and the reduced robot's speed. In this context the masterpieces of the exhibition were very close to each other, and when the robot began to move slowly compared to the visitors speed was evident the next position of the robot; so the groups were often ahead to that position.

The identification and classification of spatial formations and proxemics behaviors are not very accurate and relies to some extend to observer's subjective assessment. Provided social spacing is a key behavior in robotic guides, a thorough description of these categories and a measure of inter-rater reliability should be addressed for further studies.

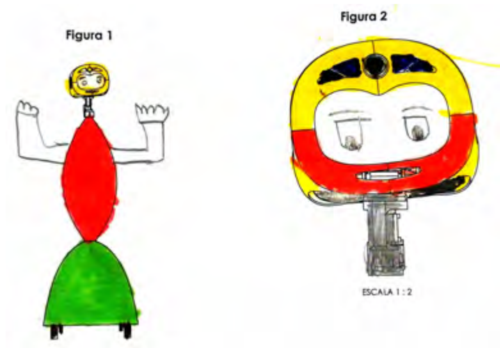


Figure 5.9: Drawing and painting of the body and head of the robot made by one of the visitors

The reasons for preferences of particular age groups and why interactions are generally in larger groups are questions that remain unresolved. One possibility is that there is a higher proportion of visitors with these features in this social scenario, for which a previous study of the site may be appropriate.

Although it was designed for all audiences, children mostly participated in this HRI experience, so the result of the survey is biased. For instance, it makes sense that they found the robot to be tall. An implication in the design could be that the robot can vary its height depending on the age of the user.

Depending on the context, on the feedback, if people understand what the robot is for, what's its role, when it is more appropriate to do one thing or the other. For example, although it is true that in its vast majority people preferred the activities of dialogue and play (75%), it should be noted that in this special context the largest number of users were children. In addition, it was not very explicit to differentiate the roles of the robot. It would be necessary to analyze the convenience or not that a robot can have multiple roles or the way in which the different roles of the robot are made known to the people.

On group behavior, more specifically a dyadic group-robot, it is interesting to note that most visitors like to be close to the robot and accompany the robot during the exhibition, while a smaller percentage do not like to follow or be followed by the robot. This could denote a treatment (of the participants towards the robot) more social, more egalitarian, like a peer, in the case of dyadic groups. Because the exhibition stands were quite close to each other, there was possibly no need for people to follow the robot during navigation.

One of the most accepted skills was the fact that the robot can recognize the name of the participants.

In general the design of the head, its movements and its face, had a very good acceptance.

On the characteristics of size and weight can be seen the congruence between the high percentage of people who indicated no fear of the robot, with the low percentage of participants who thought that the robot was too heavy. For a lower percentage, visitors believed that the robot was very tall.

The majority of visitors and specially children liked to be addressed by their name, and many of them spontaneously facilitate the robot personal information such as their age, in an attempt to engage in conversation and probably to explore the boundaries of robot autonomy and intelligence. This behavior point out the potential of an eventual customization of the dialogue to enhance the feeling of social awareness. However, this also puts on the table the ethical problem about the treatment and privacy of people's data,

In relation to robot's navigation some visitors mentioned that the robot was too slow. Provided displacement is one of the key skills in the role of guide, further work is to be developed to adapt robot's speed to visitors' expectations into the above mentioned safety constrains of navigation in a public crowded space.

Some additional notes about the HRI experience: in several dialogue during interaction with the robot guide, visitors -mostly children- asked to the robot about its personal information like its name, age, where was he born and even whether he had a girlfriend. This makes us reflect on the convenience or not of assigning the robot a personality, such as name, age and gender.

Due to technical problems related to the text-to-speech functionality, in the trials robot's verbal communication was based on the teleoperator's voice that speaks with the visitors in a natural, open and human-like dialogue through the robots loudspeakers. We consider that this situation interferes the WOZ technique and what is seen more like an inter-mediator with a remote somebody else than a social agent itself. To enhance the illusion of autonomy we are considering the use of a more robot-like language using synthesized voice and a limited predetermined repertoire of words and simple sentences according to a script. In addition, this script dialogue modality of communication would make the analyses of the interaction episodes easier by reducing the variability of the dialogues.

There were situations in which due to the noise of the environment or the spatial arrangement of the visitors around the robot, it was not possible to listen to what the visitors were talking about, but also the visitors could not hear what the robot said. Therefore, to make the HRI more social and effective, it is important to take into account the volume control of the voice and sounds of the robot depending on the voice volume of the people, the spatial arrangement and the noise of the environment.

5.1.5 Conclusion

An exploratory study on group-robot interaction in the context of a Cultural Center was carried out in order to observe visitor's preference, their characteristics and their behaviors.

The robot succeeded in developing roles as an exhibition guide, playing with people and maintaining dialogues, using wizard-of-oz technique. 32 interactions were observed and analyzed. The analysis was focused on visitor's as a groups more than as an individual. Groups were described according to their age and size, while the behavior were analyzed in terms of f-formations and proxemic behavior. Observational methods applied to evaluate group-robot interaction provide fruitful insight to understand the group-robot interaction by means of human-robot spatial relationships.

On the other hand, the survey conducted, although moderate in number, yielded interesting results that will help us in future designs of the robot and the HRI in general.

5.2 MASHI at “Vaixells a la mar” Exhibition

In this section we describe a second HRI field trial with our MASHI robot as an tour guide in the exhibition “Vaixells a la mar” (*Boats at sea*)(Figure B.4) at the Bóbila Cultural Center. In this study, in addition to testing the improvements performed to both, the robot and the HRI, from the first experience, we want also to measure the users' perception of robots through a survey.

This section is organized as follows: In Subsection 5.2.1 the study design is detailed. The results of the survey carried out are presented in the Subsection 5.2.3. In Subsection 5.2.4, discussion and conclusion from the results obtained from this experience are presented.

5.2.1 Study Design

This section discusses general issues related to the design and development of this second experience in La Bòbila. The MASHI robot was deployed in the main hall ($6 \times 8m^2$ aprox.) at La Bòbila Cultural Center. The field trial was carried out for one week, for about 2 hours per day at the afternoon.

Objective

The main aim in this experience is to observe social human-robot interactions *in-the-wild* with MASHI as an exhibition guide. In this experience our specific objective is to measure the users's perception of MASHI from this experience.

Scenario and setup

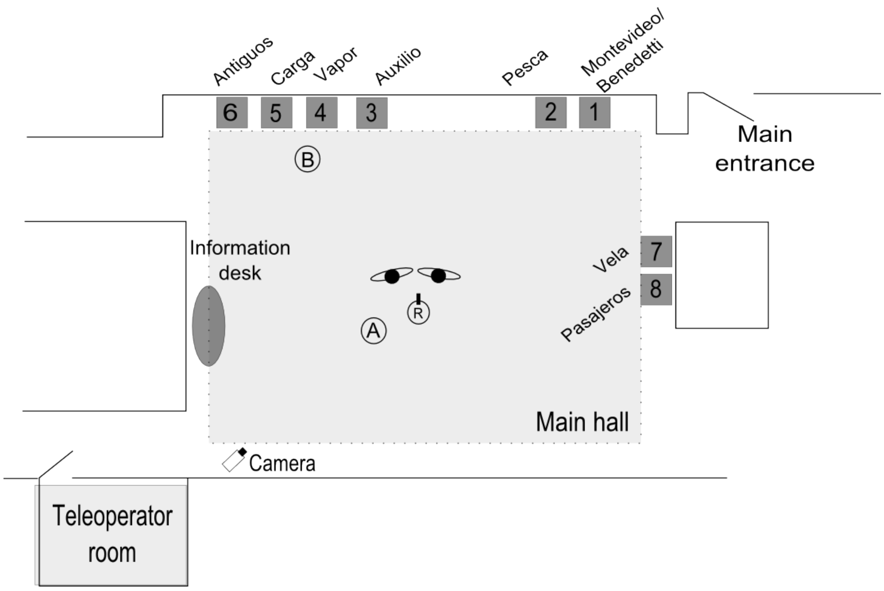
Like the previous one, this HRI experience took place in the main hall of the Bóbila Cultural Center. Two locations (A and B) were defined as the main possible destinations for the robot guide. Point A is the initial location of the robot in the center of the hall and in front of the information desk. Point B represents the exhibition guiding area, comprising eight works of the exhibition (see Figure 5.10b).

The field study was carried out from one week for about 2 hours per day at the evenings coinciding with the hours of greatest concurrence to the Cultural Center. No adaptation of the physical environment was implemented to maximize the study ecological validity preserving the natural every-day conditions and routines except from a zenital-like camera with a fisheye lens placed in the second floor (see Figure 5.10a) out of sight of visitors at a height of approximately 3 meters in order to have an overall view of the scene (Figure 5.10b)

The robot

MASHI v0.4 (see Figure 5.11) is the improved experimental robotic platform used for this study of social human-robot interaction. With a mobile base and lightweight structure, the robot is $1.5m$ tall and weighs about $15.0Kg$. The upper part of the robot

(a)



(b)

Figure 5.10: "Vaixells a la mar" scenario and exhibition setup.

comprises a torso and a motorized head with yaw, pitch and roll movements. The front of the head features a 7" inches wide angle display that serves to show an animated face (i.e. eyebrows, eyes and mouth) to support non-verbal communication by its facial expressions. To avoid some type of bias due to the appearance of the face of the robot, only the features that help generate facial expressions, such as eyebrows, eyes and mouth have been left. At the torso level the robot has a stereo speaker and a microphone. The mobile base endows 2 degrees of freedom, with two powered wheels and two caster wheels for its stability. In this study, MASHI robot attempts to move at $0.16m/s$ and seeks to turn at $0.74rad/s$. In this experience the robot has a camera with a fish-eye lens just above his face to give a panoramic front view. MASHI v0.4 has a kind of clavicle, from which ends two arms are detached. Each arm has two degrees of freedom: one active or motorized at the level of the elbow and the other passive or non-motorized at shoulder level.

In the operator's side, the teleoperation system is developed under the WebRTC platform, which allows a full-duplex real-time communication of audio, video and data. The operator could move the robot base back and forth and rotate left or right, make pitch, yaw and roll head movements, and play music, using the keyboard or buttons in the computer interface. The teleoperator room was just next the main hall in a private room inaccessible for visitors (Figure 5.10b).

Task

The robotic platform was used in a Wizard-of-Oz setup thus the robot's head movements, displacements, dialogue with visitors (through text-to-speech) and interactions were totally teleoperated by the operator that remains out of visitor's sight inside the operator room.

To make a more enriching HRI experience, a script (view Figure B.5) has been made as a flowchart (see Figure 5.12). The robot is initially looking for some person or group of people (we will call it a target), when it finds the target, MASHI greets, presents itself and asks if they want to take a tour. If it is affirmative, the robot goes to the initial position and starts the exposure. The exhibition contains an introduction, and the explanation of each of the bases; After which he asks the visitors if they want to know more about the author. In case of a favorable response, MASHI provides some additional information or curiosities about the work and its author the artist, Mr. Tucho

Figure 5.11: MASHI v0.4 used in this study.

Bergeret. Then, MASHI proceeds to ask the visitors if they can fill out a questionnaire, and if the answer is yes, the robot indicates where the questionnaire is (information desk) (see it in Figure5.10b) and then start the guidance cycle again.

Regarding the behavior of the robot in terms of the HRI, the generation of non-verbal language, such as the movement of the head, facial expressions, etc. they are at the discretion of the operator. Aware that the problem with not structuring the type of responses is that the robot is biased on the operator's own mood, we want the robot to be able to use the greatest amount of resources to communicate with people.

No briefing or instruction was given to visitors, and the intervention of technical staff at the local environment was exclusively aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

5.2.2 Data Analysis

To measure the users' perception of MASHI during this experience, we will use the Godspeed Questionnaire Series [32], which evaluates different aspects of the robot, such as: Anthropomorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety. Each aspect of the robot contains a questionnaire that uses semantic differential scales

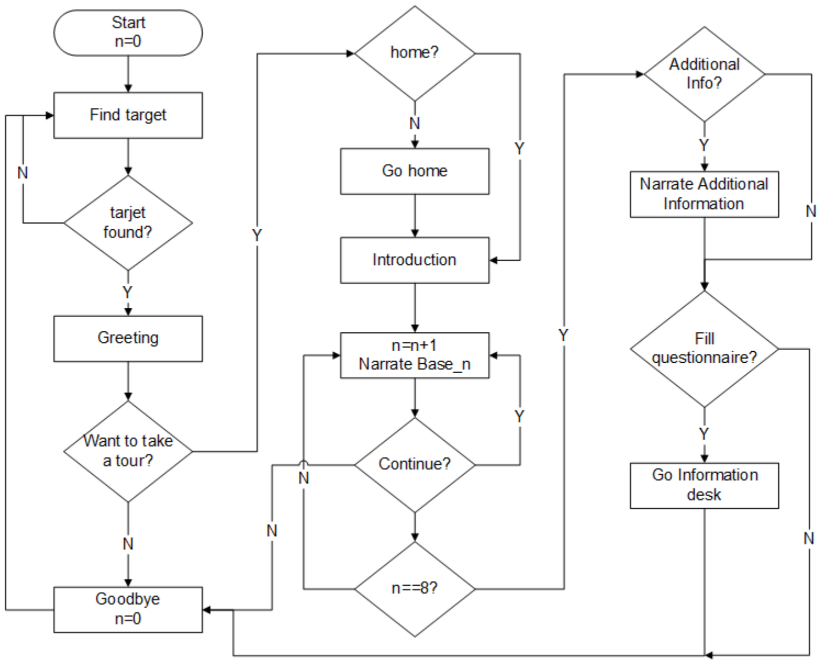


Figure 5.12: Flowchart of robot’s role in ‘Vaixells a la mar” exhibition.

with a Likert scale from 1 to 5. The scale of value 3 represents a neutral scale, so only scales 4 and 5 will be considered for positive evaluations and 1 and 2 for negative evaluations.

A manera de referencia, se

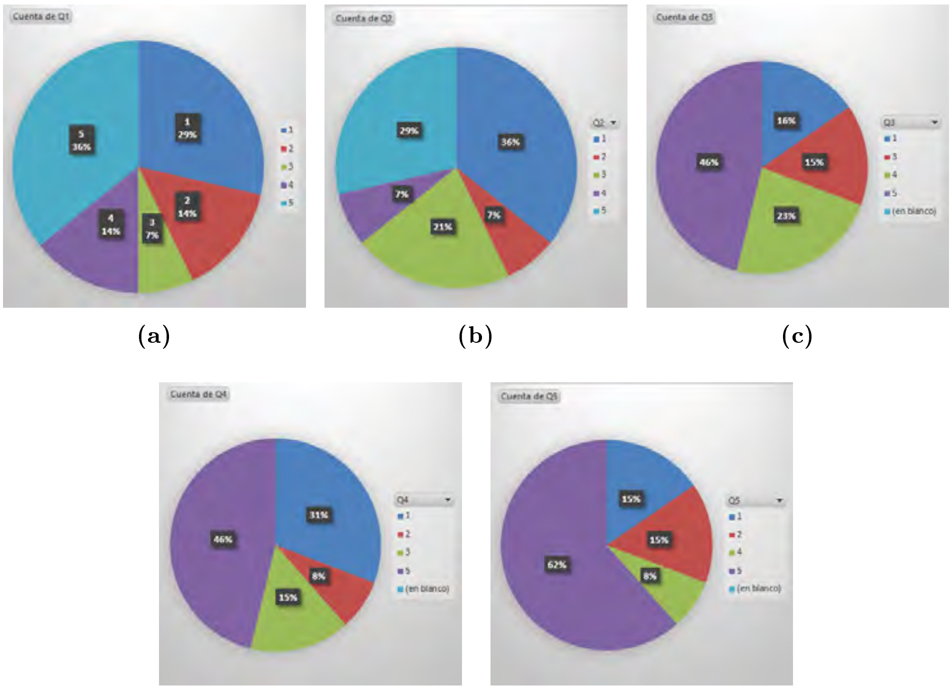
In order to measure the reliability of the different questionnaires we use the criteria of Cronbach’s alpha [49]: Excellent (alpha>0.9), Good (0.7<alpha<0.9), Acceptable (0.6<alpha<0.7), Poor (0.5<alpha<0.6), Unacceptable (alpha<0.5).

As seen in the task assigned to the robot, part of the task was to invite the participants to fill out the survey at the end of the exhibition.

5.2.3 Results

A total of 12 visitors conducted the survey, with an average age of 18 years and a standard deviation of 15, whose younger age was 9 years while the oldest age was 63 years; of which 36% are males while 64% are females.

The Cronbach’s alpha coefficients obtained to measure the reliability of the scale were: 0.78 for Anthropomorphism, 0.75 for Animacy, 0.99 for Likeability, 1.0 for Per-



(a) (b) (c) (d) (e)

Figure 5.13: Survey results of robot Anthropomorphism: (a) Q1. Fake(1) - Natural(5); (b) Q2. Machinelike(1) - Humanlike(5); (c) Q3. Unconscious(1) - Conscious(5); (d) Q4. Artificial(1) - Lifelike(5); (e) Q5. Moving rigidly(1) - Moving elegantly(5).

ceived Intelligence and 0.93 for Perceived Safety.

Anthropomorphism As we can see in Figure 5.13, the results reveal the perceptions about the anthropomorphism of the robot: 50% of interviewed people think that the robot is natural against 43% that it is not (Figure 5.13a); 36% of them think that the robot has a human aspect versus 43% that thinks it looks like a machine (Figure 5.13b); 46% of people think that the robot is conscious while 31% think it is not (Figure 5.13c); moreover, up to 46% think that the robot seems alive against a 39% that thinks it is artificial (Figure 5.13d); finally 62% is the percentage of people thinking that the robot moves elegantly against 30% expressing that it moves rigidly (Figure 5.13e).

Animacy As we can see in Figure 5.14, the results reveal the perceptions about the animacy of the robot: 38% of people declare that the robot is alive against 31% ex-

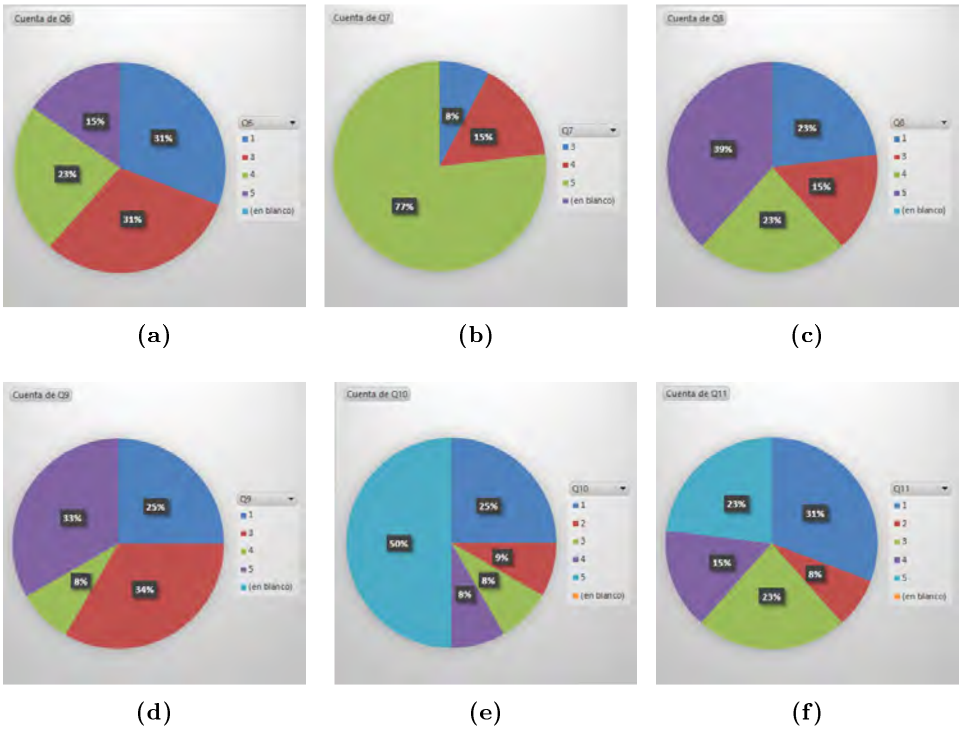


Figure 5.14: Survey results of robot Animacy: (a) Q6. Dead(1) - Alive(5); (b) Q7. Stagnant(1) - Lively(5); (c) Q8. Mechanical(1) - Organic(5); (d) Q9. Artificial(1) - Lifelike(5); (e) Q10. Inert(1) - Interactive(5); (f) Q11. Apathetic(1) - Responsive(5).

pressing it is dead (Figure 5.14a); most of them, 92%, think that the robot is lively (Figure 5.14b); up to 62% think that the robot is organic while 23% think it is mechanical (Figure 5.14c); only 41% of people express that the robot if lifelike against a 25% that thinks it is artificial (Figure 5.14d); more than a half, 58%, declare that the robot is interactive against 34% expressing it is inert (Figure 5.14e); finally 38% observe the robot as responsive against 39% that thinks it is apathetic (Figure 5.14f).

Likeability In Figure 5.15 are depicted the results revealing the perceptions about the likeability of the robot: almost all the participants, 92%, likes the robot (Figure 5.15a); 72% of them declare that the robot is friendly while 7% define it as unfriendly (Figure 5.15b); 79% of participants opine that the robot is kind while it is unkind for 21% (Figure 5.15c); 69% express that the robot is pleasant against 8% that vote for it to be unpleasant (Figure 5.15d); finally 62% of the interviews reflect that the robot is nice

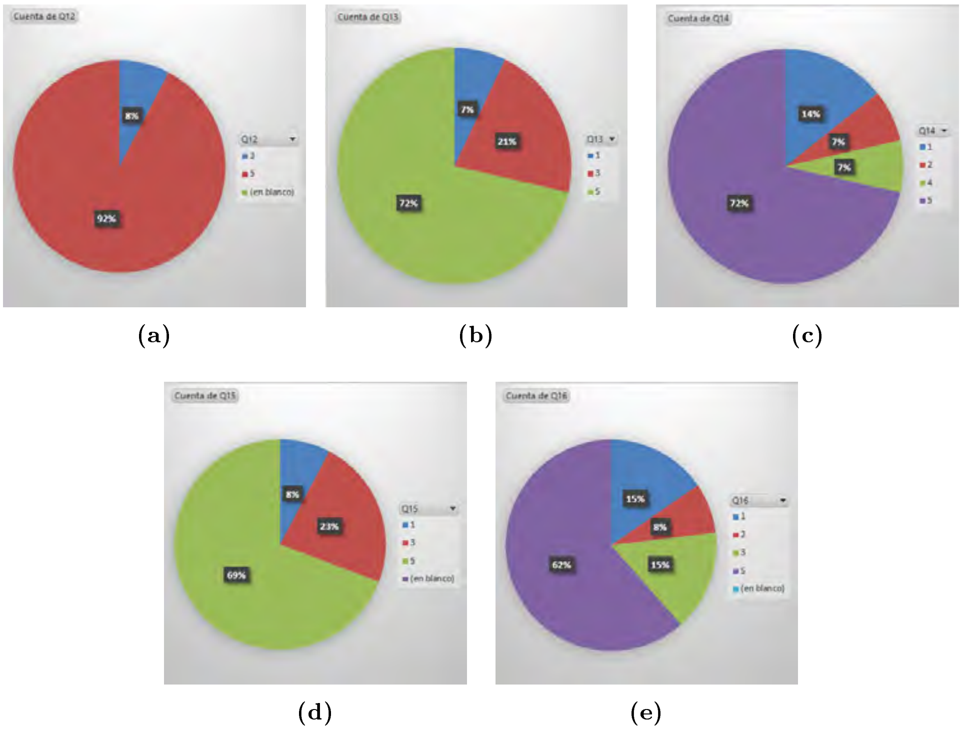


Figure 5.15: Survey result of robot Likeability: (a) Q12. Dislike(1) - Like(5); (b) Q13. Unfriendly(1) - Friendly(5); (c) Q14. Unkind(1) - Kind(5); (d) Q15. Unpleasant(1) - Pleasant(5); (e) Q16. Awful(1) - Nice(5).

against 23% that would show it as awful (Figure 5.15e).

Perceived Intelligence As we can see in Figure 5.16, the results reveal the perceptions about the perceived intelligence of the robot: 67% think that the robot is competent while 25% think that is incompetent (Figure 5.16a), 85% think that the robot is knowledgeable while 15% think that is ignorant (Figure 5.16b), 92% think that the robot is responsible while only 8% think it is irresponsible (Figure 5.16c), 92% think that the robot is intelligent while 8% think it is unintelligent (Figure 5.16d) and 82% think that the robot is sensible (Figure 5.16e).

Perceived Safety For this feature, as it can be seen in Figure 5.17, results reveal the perceptions about the likeability of the robot: up to 62% of the participants declare that the robot is relaxed while 15% think is not (Figure 5.17a); almost all of the people,

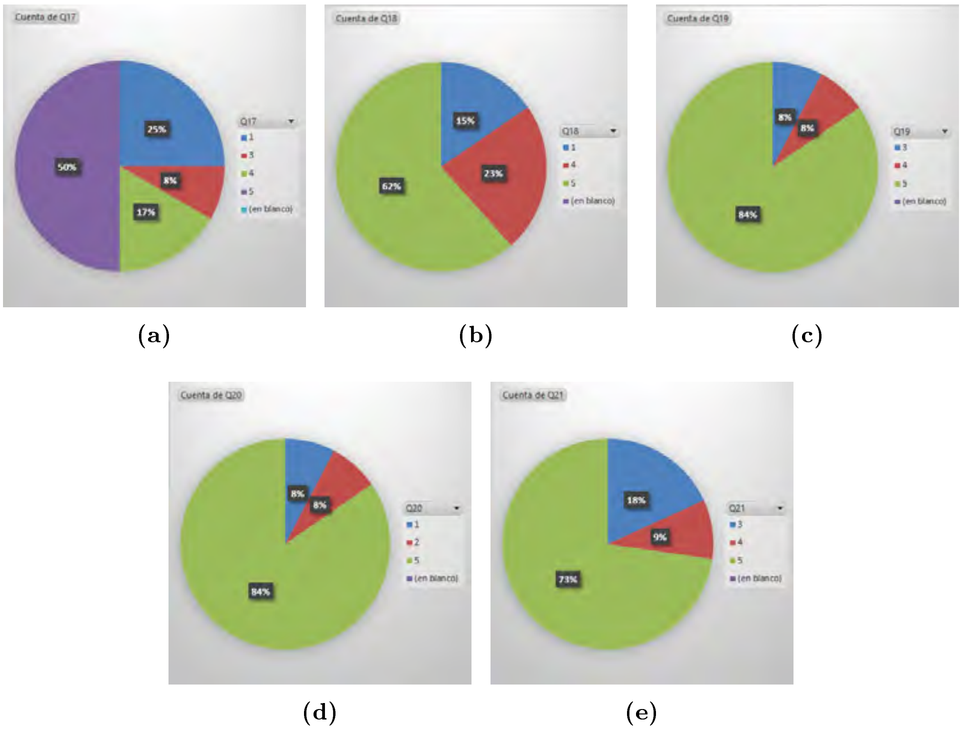


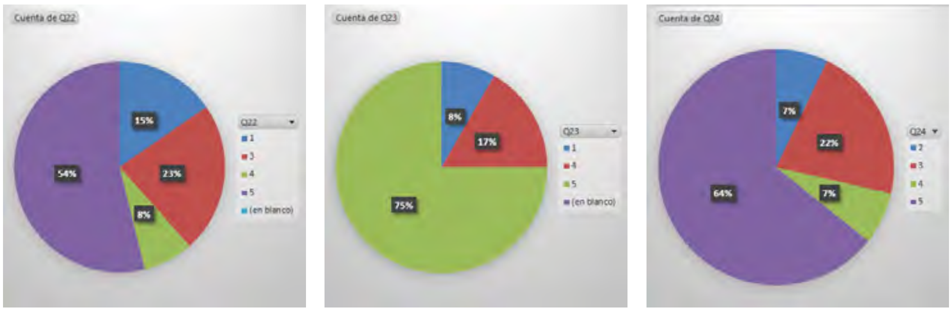
Figure 5.16: Survey results of Perceived Intelligence: (a) Q17. Incompetent(1) - Competent(5); (b) Q18. Ignorant(1) - Knowledgeable(5); (c) Q19. Irresponsible(1) - Responsible(5); (d) Q20. Unintelligent(1) - Intelligent(5); (e) Q21. Foolish(1) - Sensible(5).

92%, express that the robot is calm while 8% perceive it is agitated (Figure 5.17b); finally, 71% of interviewed people note that the robot is surprised while only 7% think it is quiescent (Figure 5.17c).

5.2.4 Discussion and Conclusion

In general, results obtained in the anthropomorphism section qualify the MASHI robot as natural, having consciousness, alive and moving elegantly. In spite of this, there is a greater percentage of those who consider that the robot is more similar to a machine than a human. This appreciation corresponds to the fact that what we want to enhance is the service offered by MASHI to the user, beyond of the aesthetic appearance of it.

The animacy section shows that a greater percentage of the participants think that the robot is more organic, more lifelike and more interactive. It should be noted the there exists a huge percentage of participants who see the robot as lively.



(a) (b) (c)

Figure 5.17: Survey results of Perceived Safety: (a) Q22. Axious(1) - Relaxed(5); (b) Q23. Agitated(1) - Calm(5); (c) Q24. Quiescent(1) - Surprised(5).

Talking about likeability, this feature got a good rating. The vast majority of participants liked the robot, while an important majority thinks that the robot is friendly, kind, pleasant and nice.

In perceived intelligence the robot also obtained a very good valuation in all its items; mainly in intelligence and be responsible, followed by knowledgeable, sensitive and competent.

Making a comparison, we can appreciate that the likeability and the perceived intelligence obtain a better valuation than the anthropomosphism and the animacy. This is according to our purpose, which is to have a better HRI experience, beyond the appearance of the robot. Instead of an appearance or human movements, the design of the robot has anthropomorphic proportions and the movements necessary for an effective verbal and nonverbal communication.

When studying the perceived safety there was an error in the transcription in the introduction to the questionnaire. It was written “Please rate your impression of the robot on these scales” instead of “Please rate your emotional state on these scales”. Despite this fact, we can check that there is a good valuation in all of the items. It would be interesting to be able to contrast the perceptions that participants have of the state of the robot versus the states of the participants themselves.

5.3 MASHI at "Segon Esdeveniment d'Integració Multicultural" Exhibition

The following section describes a third field study of HRI with MASHI robot at La Bòbila Cultural Center, L'Hospitalet de Llobregat, near Barcelona, Spain. The study has taken place between October and November 2015. The robot was introduced as an exhibition guide in an unconstrained setting where people reached the exhibition room and were offered a guided tour by the robot. In this occasion MASHI is tested as an exhibition guide in "Segon Esdeveniment d'Integració Multicultural" (*Second Event of Multicultural Integration*)(see Figure B.7).

Spatial relationships taking place between the robot and the visitors have been used as outcome measurements to judge the success of the Social HRI and the potential of MASHI to engage people in the visit. In particular, we use Kendon's theory of F - formations and Hall's theorization of proxemics to be taken into account as theoretical framework.

The materials through which the analysis of the study has been performed were both, videos of the sessions with the robot and a questionnaire. Videos were recorded from an upper position from that of the interaction that avoids the recognition of participants' faces, but allows the identification of spatial formations.

This section is organized as follows: In Subsection 5.3.1 we delineate a description of the study. Next, in Subsection 5.3.2, the techniques used for data collection are described. The results of both, video observations and survey carried out are presented in the Subsection 5.3.3. Finally, in Subsection 5.3.4, discussion and conclusion from the observational analysis and the survey were presented.

5.3.1 Study Design

Objective

The aim of the exploratory study carried out was to measure the potential of the robotic platform to elicit interest and engage people in the visit of the exhibition space of the cultural center.

The questions associated to this objective are the following:

- Is there a difference in the experience of the exhibition between an unguided visit and a guided tour with the robot?
- Is MASHI able to elicit interest from the people that pass through the exhibition space?
- Can we identify specific spatial formations during the interaction with MASHI that demonstrate that the robot is able to keep people in the visit throughout the whole guided tour?
- What is the perception of people about MASHI robot during this experience?

Scenario and setup

The exploratory study took place again in the exhibition space of the Cultural Centre La Bòbila (Figure 5.1a). The space is a hall 8 meters wide and 6 meters long with big openings on both sides. On the right side of the room there is the main entrance to the cultural centre, whereas on the left there is the access to the library. In the room, there are 8 exhibitors, where handcrafted objects are placed (Figure 5.1b).

The operator room is located at one corner of the exhibition space, close to the information desk. The camera is located at the second floor of the cultural centre on one corner of the room. No adaptation of the physical environment was realized.

The robot

MASHI v0.4 (see Figure 5.11) is also used for this study of social human-robot interaction. Unlike the previous experience, this time the robot was provided with a greater number of predefined phrases (see Figure B.9), not only focused on the robot's function as a guide, but also to expand the verbal language with a greater variety of possible answers. In addition, a text box for text-to-speech was added to the operator's interface, just in case the operator wanted to write something different to the pre-established phrases.

In terms of non-verbal language, the robot can perform, through the operator, some types of behavior such as movement of the head, movement of arms to point objects or to point to the information desk where the questionnaires are located. Likewise, the speed of the mobile base can be graded in case the robot has to break through people.

Participants

The participants of the study were the daily visitors of the Cultural Centre La Bòbila. Participants were not selected through specific exclusion and inclusion criteria and belong to a wide range of age groups, from children to adolescents and adults.

Since there was no invasive video-recording and facial features and identities were impossible to be recognized from video recordings, participants were not asked to sign an informed consent. Given the particular type of recruiting of the participants, we define as participants only people that initiated a guided tour with the robot. Remaining visitors that were not involved in the tour throughout the exhibition were taken into account just for the qualitative analysis of the study.

Procedure

The robot was teleoperated remotely in a Wizard-of-Oz set up and was supposed to follow a specific script (view Figure B.8) and dialogue (see Figure B.9) with a flowchart (see Figure 5.18) during the interactions with the audience. The flowchart consisted of four different stages: a waiting phase, where the robot was in a standby state; a welcome phase, once one or more visitors drew attention; a guided tour phase, when the visitors agreed to visit the exhibition; and a good-bye phase, once the visit ended.

The robot is initially looking for some person or group of people (we will call it a target). Once it finds the target, MASHI greets, introduces itself and asks whether visitors want to take a tour. When answer is affirmative, the robot goes to the initial position and starts the exposure. The narrative of the exhibition from the robot contains an introduction and the explanation for each one of the exhibitors (see Figure B.8). When this phase is completed, MASHI proceeds to ask the visitors if they can fill out a questionnaire. If the answer is affirmative, then the robot indicates visitors where the questionnaire is (information desk) (see it in Figure5.10b) and start the guidance cycle again.

No briefing or instruction was given to visitors, and the intervention of technical staff at the local environment was exclusively aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

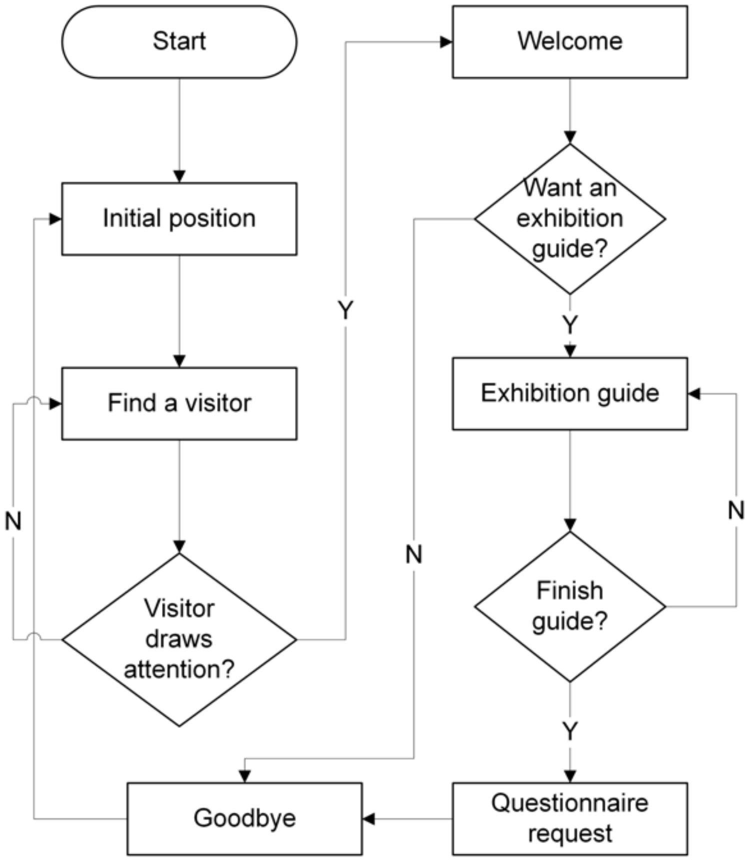


Figure 5.18: Flowchart of robot's role in "Segon Esdeveniment d' Integracio Multicultural" exhibition.

5.3.2 Data Analysis

The method chosen for the study is associated to "in field observations". Five sessions of observations have been carried out without the robot and other five have been undertaken with it. In the sessions without the robot, no guide was provided to the public. In the sessions with the robot, MASHI tried to draw the attention of people passing by the exhibition space and initiated a guided tour when the feedback from the audience was positive. All the sessions with the robot were video-recorded with a camera located on an upper floor from that of the exhibition, allowing for the recognition of F-formations and distancing patterns.

To make meaning of the interactions with the robot, we used as outcome measurements the occurrences of Kendon's F-formations. We first distinguished between guided tours and naïve interactions, then we counted the occurrences of the different f-formations within the former, and analyze the spontaneous behaviour within the latter.

To count the occurrences of the different F-formations, we developed a check-sheet (see Figure B.10) in which the observer annotated the most frequent F-formations within each minute of interaction with the robot. We performed two different types of analysis for each session. For session 6, we counted the number of the guiding service offers proposed by the robot and the number of them being accepted. For session 10, we counted the occurrences of the different F-formations during the guided tour using the check-sheet. From both videos, we annotated several interesting episodes that account for important features of HRI and reported some patterns of behaviours for what concerns proxemics. From all this material, we have been analyzing two videos with the robot (Session 6 and 10).

Questionnaires to be filled in voluntarily were left at the information desk of the Cultural Centre La Bòbila. During sessions without the robot, the questionnaire was filled out just by people stopping at the counter of the information desk, whereas during the sessions with the robot, visitors were explicitly asked to fill in the questionnaire by the robot itself. The questionnaires were two, one was a survey related to the visit of the museum (see Figure B.11), which was given during both sessions without and with the robot. The second questionnaire regarded the perception of anthropomorphism and animacy of the robot, its likeability, perceived intelligence and safety (see Figure B.12); it was provided just to participants of the sessions with the robot.

In order to establish the reliability of the questionnaires we use the criteria of Cronbach's alpha [49]: Excellent ($\alpha > 0.9$), Good ($0.7 < \alpha < 0.9$), Acceptable ($0.6 < \alpha < 0.7$), Poor ($0.5 < \alpha < 0.6$), Unacceptable ($\alpha < 0.5$).

5.3.3 Results

From the video analysis

The recorded videos we were given have a length of 2 hours (session 6) and 2 hours and 45 minutes (session 10). Videos of the HCI interface used by the operator were recorded for sessions 8, 9, 10. No video from the onboard cameras was recorded. Sessions 6 and 10 were selected because they were the ones that had the most HRI during the days.

In session 6, 4 out of 7 offers of a guided tour were accepted. Among the 4 tours accepted, 1 was not completed. The first 2 accepted offers (minute 5:58 -11:46; minute 25:09 - 32:00) seemed to work very well. Within these interactions, the F-formation created was semi-circular. In the second accepted offer, since the group was very big (11-13 people), we noted that the semi-circular configuration was repeated in several rows. To be highlighted is that the robot was not comprised in the semi-circular formation, but it was placed in front of it.

The semi-circular arrangement was created while the robot was speaking in front of the glass case. When the robot stopped speaking and started moving to another glass case, the configuration got broken to let the robot pass, and got recomposed once the robot reached the new position. It is important to emphasize the fact that in these two episodes, the robot verbalized its displacement. In other situations, when it did not do so, people could hardly understand the robot's intentions, and appeared very confused. At the end of these first two accepted tours, people spontaneously applauded the robot, a feedback that accounts very well for the quality of the interaction.

From minute 32:30 to 40:00, we assisted to a very long photo session with the robot and from minute 40:00 to 1:00:00 the robot was non moving at all, because the operator was taking a break.

At minute 1:05:23 to 1:06:50 the robot proposes the guiding service to three children. They seemed very interested to the robot, touched it and got particularly close, getting inside the intimate space. Nevertheless, they went away once the guide started. This pattern of behaviour for children appears several times in the videos. For example, at 1:16:57, where a child approaches the robot, gets very close to it at an intimate distance, touches the robot, but then leaves when the guide starts. In session 10, such behavioural pattern appears again: for instance, at minute 26:35, minute 27:15, minute 35:32. MASHI is able to draw children's attention and this is clear from their spontaneous interactions with it (minute 1:41:16), they touch the robot, shake its hand, hug it, explore it, for example waving in front of its face, nevertheless the robot does not behave in a way that they found engaging.

Another important event within session 6 appears between times 1:33:06 and 1:43:00 when a group of children plays in a very invasive way with the robot, for example shaking its hand way too strongly or putting themselves in the trajectory of the robot, preventing it to displace. Such situations make it evident that children want to understand what is the robot for and how it is supposed to behave, but could not figure it out. Moreover, it brings out the fact that the role of the robot is not understood. Children behave in a so intrusive way in symmetrical interactions and a guide is supposed to take the stance of a leader, therefore to have the control of the situation and act in an asymmetrical way.

For what concerns session 10, we analyzed the first 2 hours of video to understand which Kendon's spatial configurations were the most widespread (see Table 5.5). We identified 9 accepted guiding services and just 4 out of them got to their conclusion. Especially during session 10, it was really difficult to understand when the robot was offering a guide because quite often the flowchart was not followed. Furthermore, sometimes people approached the robot while it was still speaking after somebody left from a previous guide. This situation was particularly difficult for us to classify.

Table 5.5: Most widespread F-formations

	Min 1	**Min 2**	**Min 3**	**Min 4**	**Min 5**
guide 1 **8:40** **adults**	semi-circular (group)	semi-circular (group)	none	side by side (couple) concluded	
guide 2 **14:54** **adults**	none (couple)	none (couple)	none (couple)	semi-circular (couple with the robot) concluded	
guide 3 **26:35** **children**	semi-circular (couple)				
guide 4 **27:15** **children**	none (group)	none (group)			
guide 5 **35:32** **child**	none (individual)	none (individual)			
guide 6 **57:00** **mixed**	none (group)	none (group)	none (group)	none (group)	none (group) *robot is static
guide 7 **1:14:00** **child**	L-shape (individual)	none (individual)	none (individual)	none (individual) concluded	
guide 8 **1:19:00** **mixed**	none (group)	none (group)	none (group)	side by side (group)	none (group) concluded
guide 9 **1:26:15** **mixed**	side by side (couple)	none (couple)	none (couple)	semi-circular (couple with the robot)	

Within the 9 guided tours accepted, it was very scarce to identify Kendon's configuration. Usually, the F-formations appeared when the robot interacted with adults, and in this case, the semi-circular formation was the most likely to appear. From 31 minutes of interaction, 5 minutes were spent in a semi-circular shape, 3' in a side by side shape, 1' in an L-shape, and 22' in none of the F-formations.

At minute 48:20 in the session 10, a huge group of people is waiting for MASHI to initiate an interaction, but the robot does not react, so little by little people leave. It is not clear if there was a safety or a technological issue going on in this occasion.

In order to produce a more intriguing interaction with children or adults, it is very interesting to pay attention to the type of conversation children try to have with MASHI:

Child 1: *Hola!*

Child 2: *Hola!*

MASHI: (delay) *Hola, buenas noches!*

Child 1: *Me llamo Llúcia, ¿cómo te llamas?*

MASHI: *Me llamo MASHI... muchas gracias, soy un robot guía de exposición...*

After a few minutes, the child asks to MASHI how old it is. Children here are looking for MASHI's story. It is for them a plausible character physically, about whom they want to know more. Nevertheless, even though they are asking, they receive very few information. In this sense, it could be really useful to develop a character around the robot, with an age, a family, a past that could be consistent with the topic of the exhibition and could justify its presence there. The role of the serious guide interested just in speaking about the exhibition without satisfying any curiosity could be good to engage adults, but it is not enough to engage children. What is more, spontaneous interactions underline the affordances of the robots. For instance, MASHI's hands could be shaken and his torso could be hugged and all these element could be introduced in the script of the interaction.

Another issue to be solved is displacement, it is really hard for visitors to understand what will happen next, where is MASHI going when it finishes an explanation. This is a problem that could be easily solved through a very well written script that makes reference to the environment where the displacement takes place: for instance, to the number of the glass case the robot is moving on to, or through the exploitation of the environment configuration, for example in the current setting, the guide could follow the succession of glass cases, instead of moving randomly from one case to the other.

From the survey

We will analyze answers provided by visitors when filling in the questionnaires.

Without the robot A total of 8 people answered the questionnaire. Among the 8 surveys, only 4 of them indicated age and gender, resulting in an average age of 49 years and a standard deviation of 10; while 3 people indicated being male and one female.

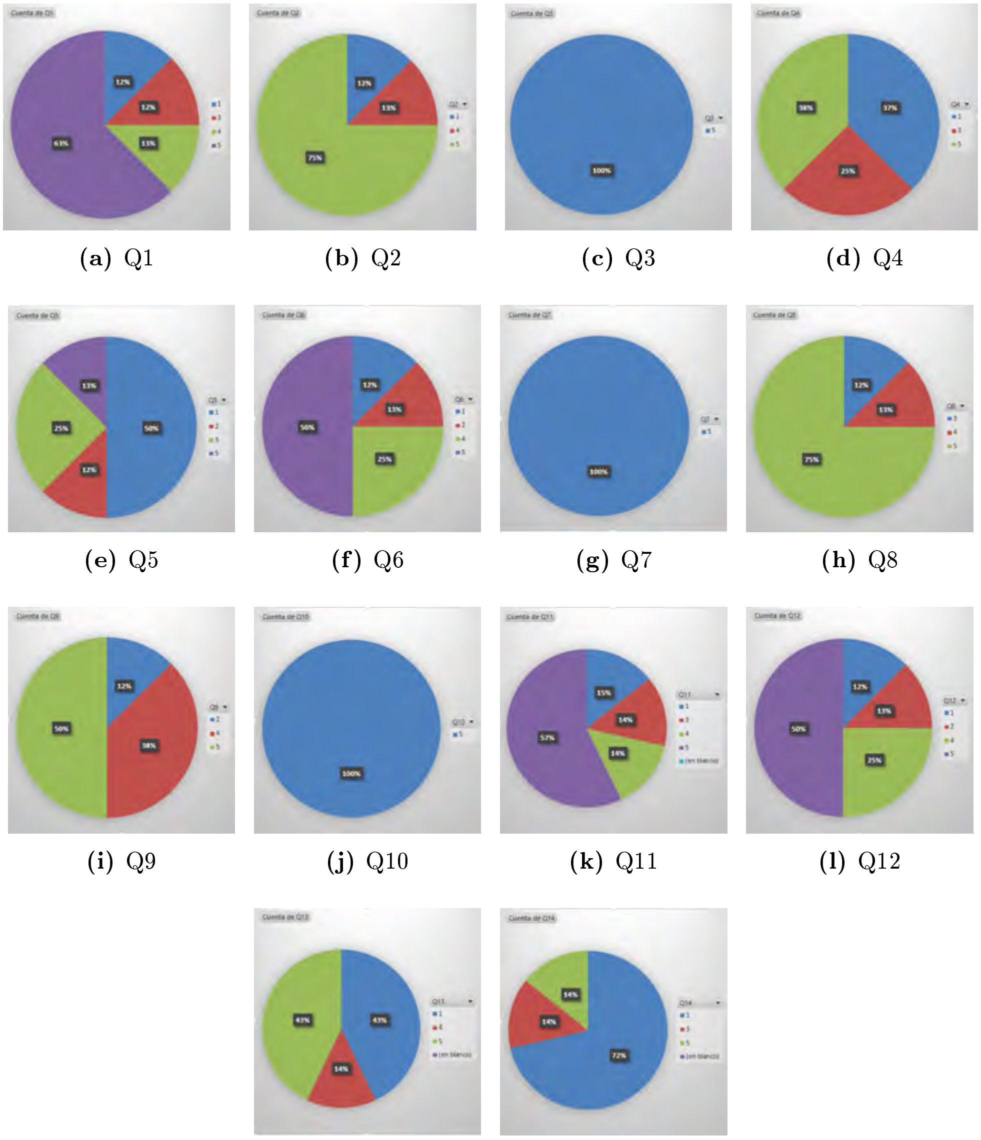
The Cronbach's alpha found in the first 14 questions of the questionnaire (in the Likert scale) was 0.73 (good).

In Figure 5.19 are depicted the results (Q1 to Q14) about the exhibition in general, without the robot. In the last questions (Q15 to Q18), for most of the visitors (87%) it was the first time he visited the exhibition (Q15); 75% visited all 8 showcases, 13% visited 7 showcases and 12% visit 3 showcases (Q16). Finally, 14 minutes on average with a standard deviation of 9 are the minutes the participants declare there are enough to cover the entire exhibition (Q17) and on average they have been accompanied by 6 people with a deviation of 1.7 (Q18).

With the robot A total of 17 visitors answered the questionnaire, of which 16% were male and 54% were female. Average age was 69 years old with a standard deviation of 11. The Cronbach's alpha found for the first 14 questions of the questionnaire (in Likert scale) was 0.83 which is a good value.

In Figure 5.20 are depicted the results (Q1 to Q14) on the exhibition in general with the robot.

In addition we can make a comparison in the knowledge acquired in the exhibition without the robot Vs the exhibition with the robot, as shown in Figure 5.21. The questions related to the knowledge acquired are from Q8 to Q14. It can be seen that in general, the results of the guidance without the robot are slightly superior to those guided with the robot, except in question Q14, which refers to whether he knew any anecdote of the author. The results can not be considered as conclusive due to the low number of participants both in the tour without robot (8) and in the tour with robot (17)



(a) Q1 (b) Q2 (c) Q3 (d) Q4

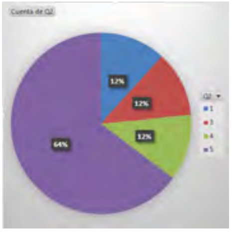
(e) Q5 (f) Q6 (g) Q7 (h) Q8

(i) Q9 (j) Q10 (k) Q11 (l) Q12

(m) Q13 (n) Q14

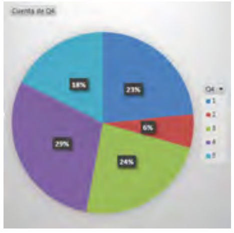
Figure 5.19: Survey results about the exhibition.

(a) Q1



(b) Q2

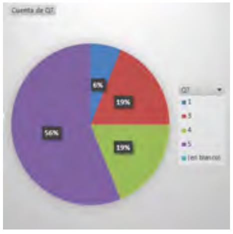
(c) Q3



(d) Q4

(e) Q5

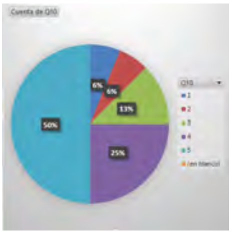
(f) Q6



(g) Q7

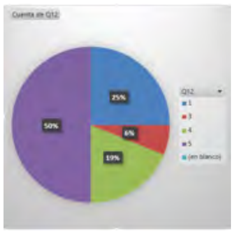
(h) Q8

(i) Q9



(j) Q10

(k) Q11



(l) Q12

(m) Q13

(n) Q14

Figure 5.20: Survey results about the exhibition with robot.

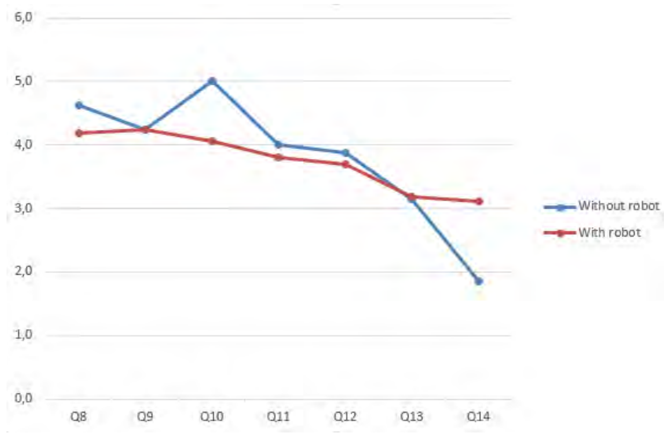


Figure 5.21: Knowledge acquired between exhibition without robot and with robot.

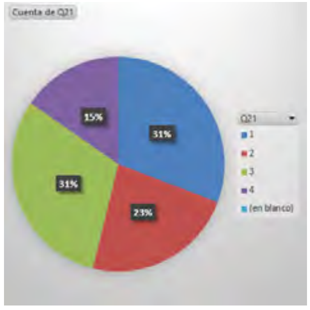
From questions Q15 to Q18, for 56% of people it was the first time they visited the exhibition (Q15) and 65% visited all 8 showcases, 29% visited 4 showcases and 6% visit one showcase (Q16). Finally, 18 minutes on average with a standard deviation of 14 are the minutes the participants believe there are enough to cover the entire exhibition (Q17) and on average they have been accompanied by 6 people with a deviation of 5 (Q18). The Cronbach's alpha coefficients obtained to measure the reliability of the robot perceptions questionnaires were: 0.84 for Anthropomorphism, 0.92 for Animacy, 0.86 for Likeability, 0.94 for Perceived Intelligence and 0.84 for Perceived Safety.

Anthropomorphism As we can see in Figure 5.22, the results reveal the perceptions about the anthropomorphism of the robot: 66% of interviewed people express that the robot is natural against 7% that it is not (Figure 5.22a); 40% of them declare that the robot has a human aspect versus 40% that writes it looks like a machine (Figure 5.22b); 15% of people highlights that the robot is conscious while 54% think it is not (Figure 5.22c); moreover, up to 36% opined that the robot seems alive against a 50% voting it is artificial (Figure 5.22d); finally 40% is the percentage of people thinking that the robot moves elegantly against 33% expressing that it moves rigidly (Figure 5.22e).

Animacy As we can see in Figure 5.23, the results reveal the perceptions about the animacy of the robot: 40% of people declare that the robot is alive against 13% expressing it is dead, while a 47% is neutral (Figure 5.23a); 53% think that the robot is lively while a 14% think is inactive (Figure 5.23b); up to 42% think that the robot is

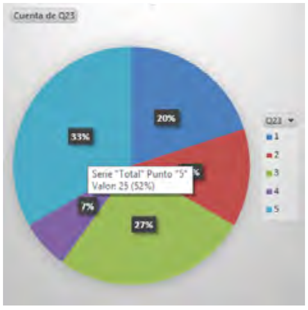
(a) Q19

(b) Q20



(c) Q21

(d) Q22



(e) Q23

Figure 5.22: Survey results of robot Anthropomorphism. (a) Q19. Fake(1) - Natural(5); (b) Q20. Machinelike(1) - Humanlike(5); (c) Q21. Unconscious(1) - Conscious(5); (d) Q22. Artificial(1) - Lifelike(5); (e) Q23. Moving rigidly(1) - Moving elegantly(5).

organic while 29% think it is mechanical (Figure 5.23c); 47% of people express that the robot if lifelike against a 40% that thinks it is artificial (Figure 5.23d); 40% declare that the robot is interactive against 27% expressing it is inert (Figure 5.23e); finally 36% observe the robot as responsive against 28% that thinks it is apathetic (Figure 5.23f).

Likeability In Figure 5.24 are depicted the results revealing the perceptions about the likeability of the robot: 80% likes the robot(Figure 5.24a); 86% of them declare that the robot is friendly (Figure 5.24b); 87% of participants opine that the robot is kind (Figure 5.24c); 86% express that the robot is pleasant (Figure 5.24d); finally 74% of the interviews reflect that the robot is nice against 6% that would show it as awful (Figure 5.24e).

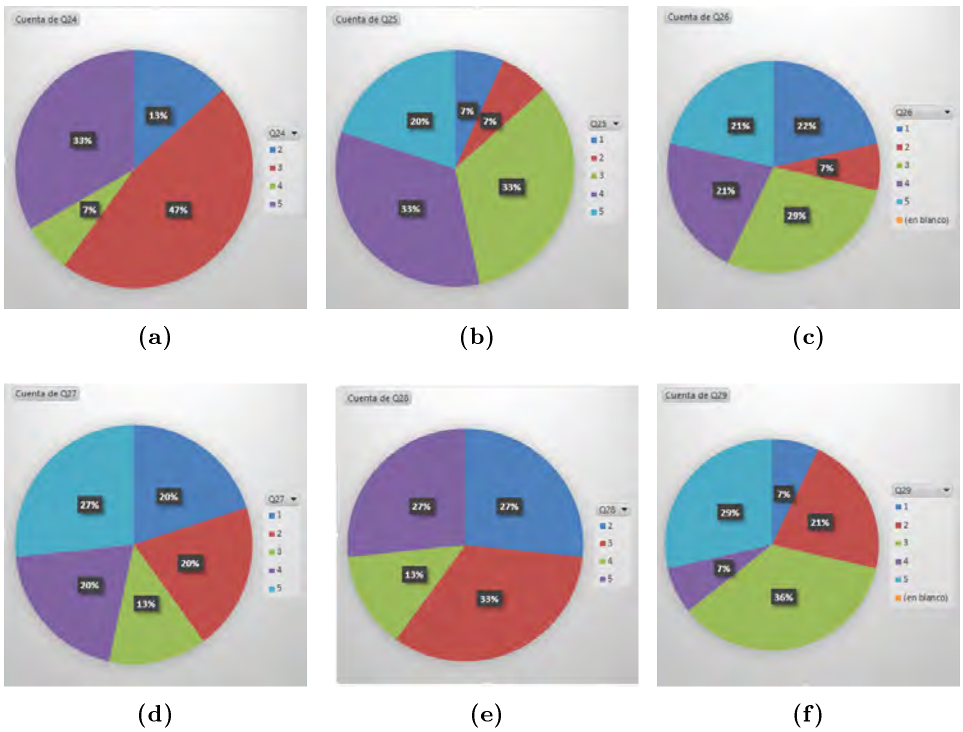
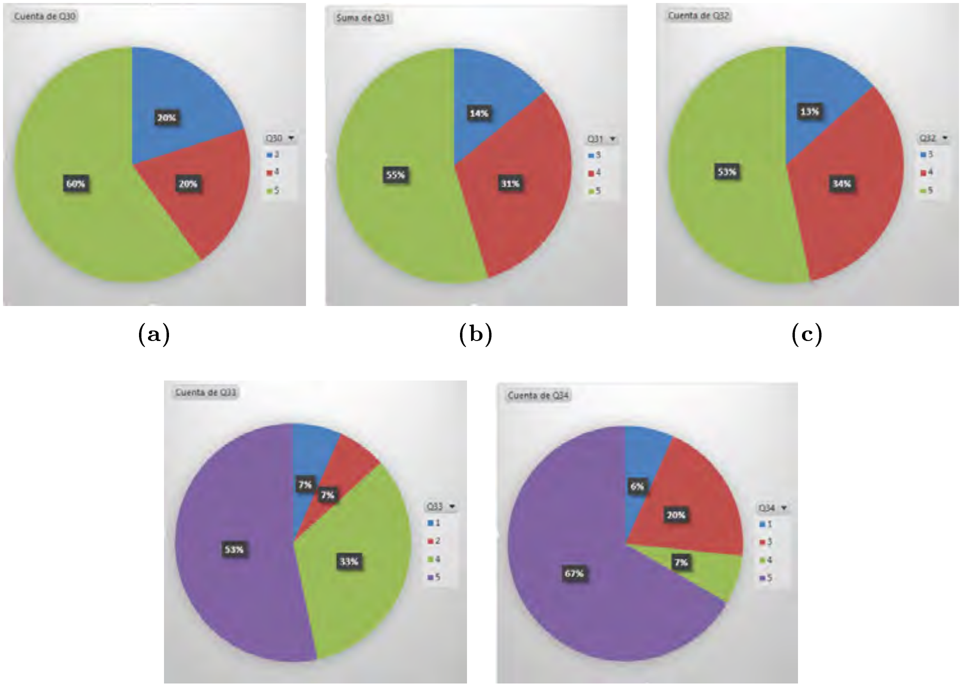


Figure 5.23: Survey results of robot Animacy: (a) Q24. Dead(1) - Alive(5); (b) Q25. Stagnant(1) - Lively(5); (c) Q26. Mechanical(1) - Organic(5); (d) Q27. Artificial(1) - Lifelike(5); (e) Q28. Inert(1) - Interactive(5); (f) Q29. Apathetic(1) - Responsive(5).

Perceived Intelligence As we can see in Figure 5.25, the results reveal the perceptions about the perceived intelligence of the robot: 60% think that the robot is competent while 20% think that is incompetent (Figure 5.25a), 53% think that the robot is knowledgeable while 14% think that is ignorant (Figure 5.25b), 54% think that the robot is responsible while 13% think it is irresponsible (Figure 5.25c), 33% think that the robot is intelligent, 20% think it is unintelligent and 47% is neutral (Figure 5.25d) and 53% think that the robot is sensible and 20% think that is foolish. (Figure 5.25e).

Perceived Safety For this feature, as it can be seen in Figure 5.26, results reveal the perceptions about the likeability of the robot: 73% of the participants declare that the robot is relaxed (Figure 5.26a); 66% express that the robot is calm while 7% perceive it is agitated (Figure 5.26b); finally, 33% of interviewed people note that the robot is surprised, 7% think it is quiescent and 60% is neutral (Figure 5.26c).



(a) (b) (c) (d) (e)

Figure 5.24: Survey results of robot Likeability: (a) Q30. Dislike(1) - Like(5); (b) Q31. Unfriendly(1) - Friendly(5); (c) Q32. Unkind(1) - Kind(5); (d) Q33. Unpleasant(1) - Pleasant(5); (e) Q34. Awful(1) - Nice(5).

5.3.4 Discussion and Conclusion

From the analysis performed, it is quite hard to find a robust and evidence-based reply to the research questions posed. MASHI is absolutely able to draw people's attention throughout different age groups. Nonetheless, it is not able to transfer this attention to the visit and it is not capable of keeping people engaged for a long time. This is partially expressed by the amount of guides accepted during session 6, just 4 out of 7, and the number of F-formations generated in the first 2 hours of session 10, just 9 minutes out of 31 of interaction.

Apart from the description of the works, there were predefined phrases that were classified between sentences, expressions of doubt, interrogatives, answers and colloquial phrases. These last ones were at the operator's discretion their use. For example, when the phrase "are we continued?" Was used to indicate the next movement of the robot, people understood better the next action the robot was going to take. So it is important

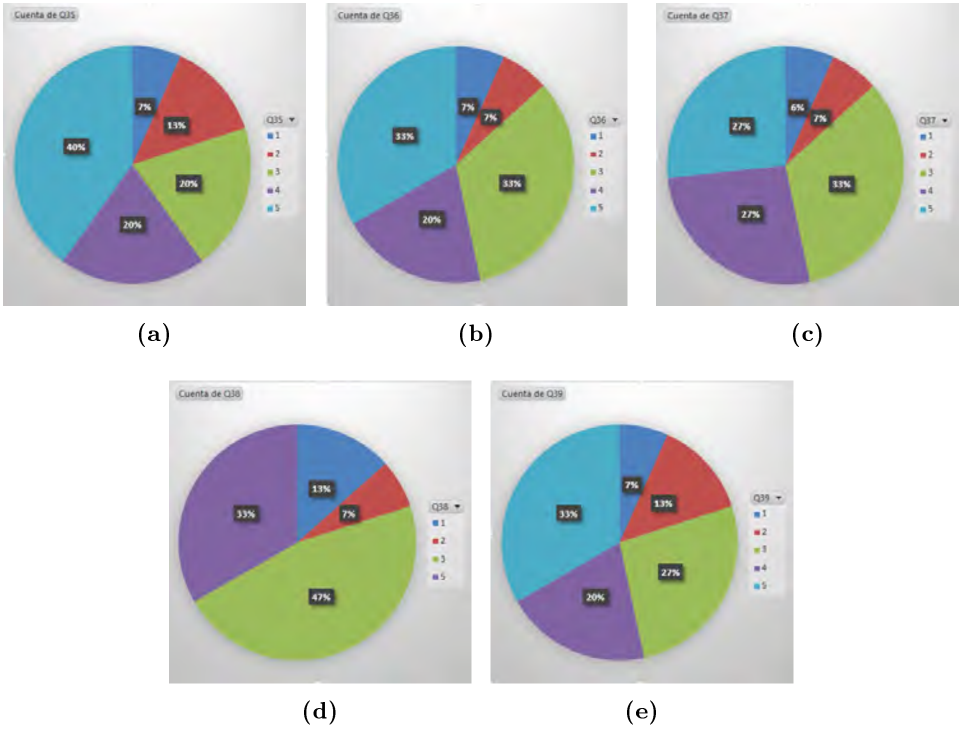
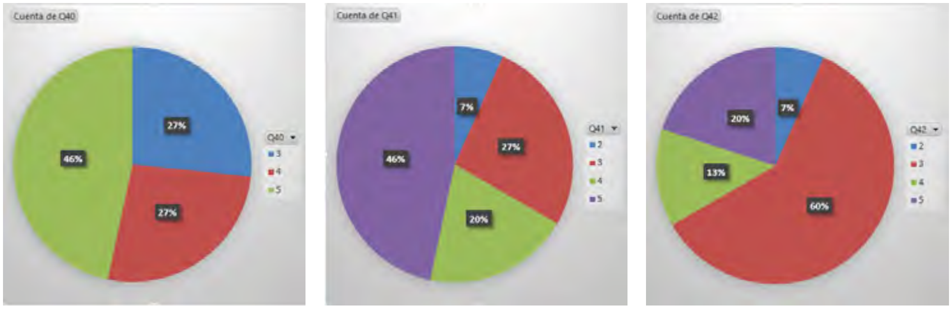


Figure 5.25: Survey results of Perceived Intelligence: (a) Q35. Incompetent(1) - Competent(5); (b) Q36. Ignorant(1) - Knowledgeable(5); (c) Q37. Irresponsible(1) - Responsible(5); (d) Q38. Unintelligent(1) - Intelligent(5); (e) Q39. Foolish(1) - Sensible(5).

to use all communication channels, both verbal (*e.g.* the use of the phrase "are we continue?") And non-verbal (*e.g.* pointing with the arm), to indicate the following actions of the robot, especially if this action implies some type of displacement.

Even though the exhibition guide was designed for all ages, it does not seem that the guide was tailored for children. However, it was observed that some children abandoned the robot during the tour, possibly because they lost interest in doing the tour. One of the factors that could have affected this was that the text or design of the HRI was not specially adapted for children, for example, making a tour more interactive and based on games and with fewer words.

MASHI elicits two different types of social HRI depending on the age group of its interactants: a more courteous and relaxed one from adults, and a very intrusive and active one from children. Such a differentiation is made evident by proxemics. Indeed, children invade MASHI's intimate space, whereas adults always maintain themselves at



(a) (b) (c)

Figure 5.26: Survey results of Perceived Safety: (a) Q40. Axious(1) - Relaxed(5); (b) Q41. Agitated(1) - Calm(5); (c) Q42. Quiescent(1) - Surprised(5).

a personal distance.

The navigation of the robot is safe, but the intentions of the robot are difficult to argue for the crowd. To solve such an issue, the robot could either use verbalization or environment affordances (following the succession of glass cases) to make displacement smoother.

The role of the robot during the guide is not clear, it should take a stronger stance towards the crowd, making its role, leading and guiding, and inner states transparent. In this last case for example, the robot could use its facial expression (anger or sadness) or a red screen colour to express its frustration when its hands are shaken too vigorously.

Conversation could be used to keep the attention of audience lively. Creating a character around a robot and presenting its story could be either a way to make children more interested to what the robot is saying. What is more, to enhance the quality of the interaction, the robot must use its body. Hand-shaking is one of the affordances that could be used during interaction.

A further study on MASHI as a guide robot should be more constrained. For instance, guided tours hours could be proposed to the visitors of the Bòbila Cultural Centre, so that the flow and number of people could be controlled and displacement could be easier. Furthermore, to understand the added value of MASHI as a guide, the situations to be compared should not be without and with the robot, but with a human guide and with a robotic one.

Last, we think that the use of F-formations as a way to measure MASHI's ability to engage people in a guided tour is not consistent. Theories more related to crowd or group behaviour and displacement could produce more significant results in such a context.

To sum up, MASHI's configuration is a very good starting point to reach a high quality social HRI, especially with children. Nonetheless, the richness of the interaction the robot is delivering should be improved and other criteria should be used to grasp the real potential of the robot in guiding and engaging people.

5.4 MASHI at the Smart City Expo World Congress

In this section we describe two experiences of the service robot MASHI in an international trade fair. MASHI is tested as a promoter robot for the "Ajuntament de L'Hospitalet" (*City Hall of L'Hospitalet*) for two consecutive years.

This section is organized as follows: In Subsection 5.4.1 we will describe the objectives of the HRI, the scenarios where it has been carried out, the participants under consideration, the characteristics of the MASHI robot that have been used, its task to develop and the data analysis that will be carried out. The results of both experiences are presented in the Subsection 5.4.2. In Subsection 5.4.3 lessons learned and the conclusions from the results obtained are presented.

5.4.1 Study Design

This section discusses general issues related to the design of the service robot and the HRI in general.

Objective

The main objective in this experience was to use the MASHI service robot as a means to attract the public, motivate their interaction and position the L'H City brand at the international trade fair "Smart City Expo World Congress".

The specific objectives are:

- Create an innovative and unique robotic service that enhance the human-robot interaction in-the-wild.
- Measure the impact of the HRI on the social network.

(a) SCEWC 2015

(b) SCEWC 2016

Figure 5.27: Scenarios at SCEWC.

Scenario and Setup

Both experiences were carried out at the "Fira Barcelona - Gran Via", with 240,000 m^2 of exhibition floor space, regarded as one of Europe's most cutting-edge venues. In particular, MASHI robot was deployed around the stand of L'Hospitalet City ($6 \times 10m^2$ aprox.) in the "Smart City Expo World Congress".

The first experience was carried out from 17th to 19th November of 2015, while a second experience was carried out from 15th November to 17th November of 2016, both for about 7 hours per day.

It is worth mentioning that in the first experience MASHI was developed only in the fair (see Figure 5.27a), while in the second experience MASHI was present on the stand accompanying the group of children and youth of the Robotics Club CortoCircuito which presented their project L ' H Smart City, a model of the city (see Figure 5.27b) product of a process of participation and co-creation (Design Thinking) for the detection of social problems of L'H and that had several social and productive actors such as the City Council of L'H, the Polytechnic Universitat Politècnica de Cataulnya, the La Bòbila Cultural Center, the "We Are Not Invisible" Association, the Affinity Foundation and the children of the Club CortoCircuito.

Participants

Most of the visitors attending the event were Executives and Professionals in public or private management, research, development and innovation in the field of Smart Cities.

(a) MASHI at SCEWC 2015

(b) MASHI at SCEWC 2016

Figure 5.28: MASHI v0.4b at SCEWC.

Robot

In these experiences, the MASHI v0.4b (see Figure 5.28) robot was used, which contained two main variants: in the robot, the left arm also had an active joint at shoulder level and a webcam on the wrist to take "selfie" photos with visitors, while on the operator side the graphical interface allows controlling the Selfie function, which consists in raising the left arm, doing a countdown, taking a photo and uploading it to social networks.

Task

In order to attract the attention of visitors, to interact with them and promote the L'H Social District brand within the fair as well as in social networks, the MashiSelfie service robot was created. Through this innovative project, MashiSelfie became the first social robot ever performing selfies and uploading photos on social networks.

The robot was teleoperated remotely in a Wizard-of-Oz set up and was supposed

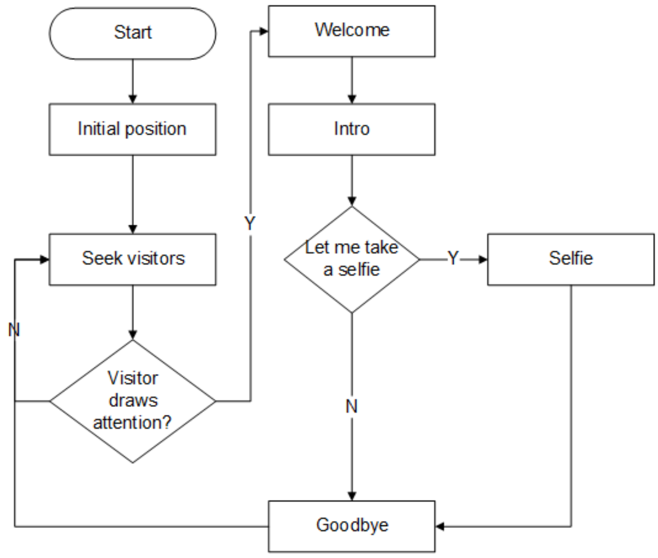


Figure 5.29: Flowchart of robot's role in SCEWC.

to follow a flowchart (see Figure 5.29) and interact with a predefined vocabulary (see Figure B.13) with the audience. The flowchart consisted of several stages: firstly, all the parameters of the system are initialized (Start); then, the robot looks for an initial position of the HRI (Initial position); next, from this position the robot keeps track of the people either, by navigating or only by moving its head ('Seek visitors' function). When people are close enough to the robot, at a personal distance, the target is chosen. When the robot keeps the attention of the visitor, a welcome message is sent (Welcome), followed by an introductory message about L'Hospitalet city (Intro); lastly, the visitor is asked whether he/she wants to take a selfie with MASHI: if the answer is affirmative a countdown is made, a photo is made and then it is sent to the twitter account of MASHI (@MashiRobot) (Selfie). Finally a message of thanks and goodbye is given (Goodbye).

No briefing or instruction was given to visitors, and the intervention of technical staff at the local environment was exclusively aimed at recovering the robot for eventual breakdowns and discouraging misuse to enhance people safety and to prevent robot's damage.

Data analysis

To measure the impact of the MashiSelfie service offered by MASHI on the social network Twitter, we will consider as relevant data information about the tweets registered during the days of the event.

In this sense, we will take two types of measures: the number of tweets generated by MashiSelfie and the number of times that this tweets has been viewed on the Twitter platform in an organic context, called organic impressions, *i.e.* the impressions generated from Tweets which does not include promoted or paid context.

5.4.2 Results

Results will be analyzed for the two SCEWC events that MASHI was attending.

SCEWC 2015

During the three days of the event, MashiSelfie managed to send 138 tweets (see Figure 5.30 and Figure 5.31), along with the selfie photo.

During this period, a total of 11,7K organic impressions were available, of which 3152 corresponded to day 1, 4761 to day 2 and 3745 to day 3 (as shown in Figure 5.32).

SCEWC 2016

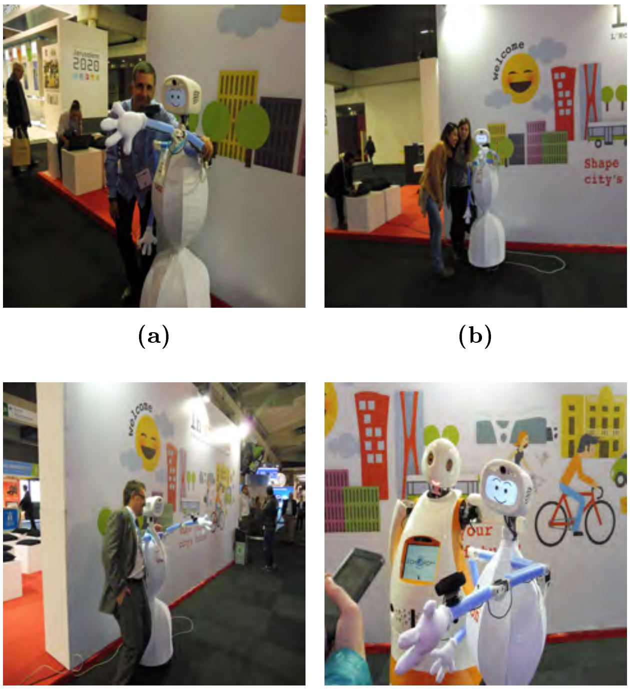
During the three days of the event, MashiSelfie managed to send 422 tweets (see Figure 5.33 and Figure 5.34), along with the selfie photo.

Moreover, during this period it was possible to have a total of 20,9K organic impressions, of which 5322 corresponded to day 1, 9746 to day 2 and 5851 to day 3 (as shown in Figure 5.35).

5.4.3 Discussion and Conclusion

It was really surprising how MASHI got people's attention. Although it was not the best robot of the Congress, nor the most modern, people from all ages were willing to take a picture with MASHI. Some people were really interested and wanted to know more about the project and the robot.

However, there were some people that didn't understand exactly the purpose of the robot and the service it was providing. Sometimes people were staying in front of the



(a) (b) (c) (d)

Figure 5.30: Several human-robot interactions and a robot-robot interaction during SCEWC 2015.

robot, and technical staff had to explain to them the task that it was performing, and how to interact with. This drawback gives an idea that's worth taking it into account in future HRI designs.

Another issue that was detected is that most of the times it is very difficult to establish a conversation between the operator and the person standing in front of the robot. The operator had to be really fast with the text-to-speech software to catch people's attention and to chat. A possible solution is that the system allows the operator to add dialogue phrases in a more dynamic way.

The metrics used are currently used to measure the impact of tweets on social networks, and are consistent with the objectives set at the beginning.

From the different dimensions of the group behavior in the natural environments that we have observed, spatial arrangements play an important role as a means to follow social norms in public environments. So far we have described the different types of spatial arrangements observed in the different interactions, but is there a way to

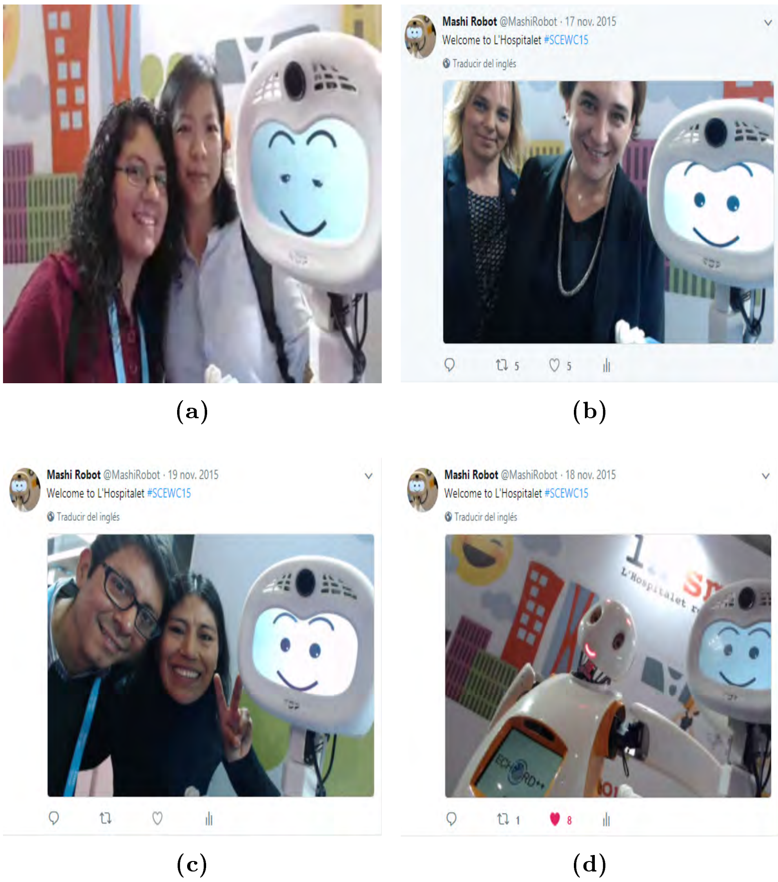


Figure 5.31: Robot selfies during SCEWC 2015 at Twitter.

characterize or describe such formations as humans do? That is, in a more intuitive or qualitative way? The next chapter is a first approximation to interpret spatial relationships in HRI using a Qualitative Spatial Reasoning - QSR approach.

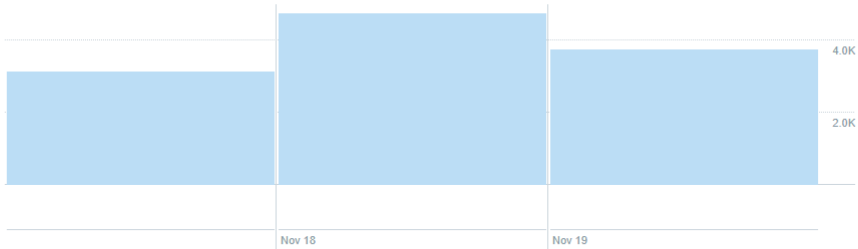


Figure 5.32: Result impressions in Tweeter from the SCEWC2015 experience.

(a) **(b)** **(c)** **(d)**

Figure 5.33: Several human-robot interactions during SCEWC 2016.

(a)

(b)

(c)

(d)

Figure 5.34: Robot selfies during SCEWC 2016 at Twitter.

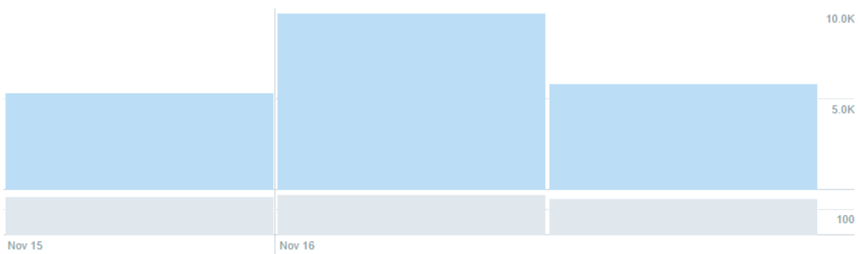


Figure 5.35: Result impressions in Tweeter from the SCEWC2016 experience.

Chapter 6

A Theoretical Analysis of Group-Robot Spatial Interaction

In this chapter we address the problem of how to represent the HR spatial relationship from a qualitative reasoning approach. In HRI two types of situations have been distinguished: fixed group, when the space occupied by the human-robot group normally does not change in the time (as in the case of a robot performer-human audience); and mobile group, that is, when the space occupied by the group changes continuously over time (as in the case of a robot guiding people).

The Qualitative Spatial model for Group Robot Interaction (QS-GRI) introduced in this chapter defines the spatial arrangements from fixed groups (i.e. F-formations) that depends on: (i) the relative location of the robot with respect to other individuals involved in the interaction; (ii) the orientation of the individuals (shared front) or not; (iii) the shared peripersonal distance (*i.e.* the distance around the body where objects can be physically manipulated); and (iv) the role of the individuals (observer, performer or interactive).

The evolution of F-formations between them must be studied: that is, how one formation is transformed into another. These transformations can depend on the role and the task that the robot have, as well as on the amount of people involved.

Qualitative approaches to represent and reasoning about moving objects have been usually defined in the literature to represent Human Robot Spatial Interactions in navigation situations where one robot and one human (or a group of humans as a whole) were involved [55][56]. Qualitative spatial representations for activity spaces where a robot carry out a task or collaborate with more that one person are not available in the literature, as far as we are concerned. This work refers to social interactions among humans (Human-Human Interaction HHI) and Human-Robot Interactions (HRI) in social environments, which may involve several individuals (sometimes arranged in group) and one robot – named as Group-Robot Interaction, GRI.

A good example of this recent interest in the community is the "Groups in Human-Robot Interaction" full day Workshop held in the IEEE International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2016)[1]. There, it is highlighted how recent studies in social psychology and HRI indicate that inter-group interaction varies crucially from inter-individual (dyadic) interaction [57]: modulating the effects found in dyadic HRI, introducing variables that are not possible to study in dyadic HRI, and requiring different technical solutions to problems of perception and interaction.

However, besides these variations between dyadic interaction and group-robot interaction, it is still interesting to find common factors that allow us to represent and to reason about the spatial relationships when interacting either with an individual or with a group. We claim that the use of qualitative metrics leading to consider the group as a whole, can help to develop techniques in group-robot spatial interaction in a more general form, allowing to inherit techniques from the usual human-robot spatial interaction. Any individual or group has a characteristic interaction region. In the case of a group, individuals often have some type of arrangement around this inner, shared region (i.e. sometimes named as *o-space* [37]). The use of this interaction region in a qualitative approach can help to represent more generally both individuals and groups in Human-Robot Spatial Interaction.

As the group HR spatial arrangements are described in a linguistic manner, we propose a qualitative model to formalize them using a qualitative representation based on distances, locations and orientation in Section 6.2. Final sections are devoted to discussion, conclusions and provide a list of future work.

[1] `https://grouprobot.wordpress.com/2017-groups-in-human-robot-interaction/`

6.1 Qualitative Spatial Representation and Reasoning Approach in HRI

From the domain of Artificial Intelligence is the research area of qualitative representation. It indicates the need and importance of representing and reasoning about the spatial aspects of the world [58].

Very few approaches in the literature have made the effort to formalize the social convention of human-robot group behavior to represent more cognitively in human populated scenarios.

In this sense, several qualitative studies use the Qualitative Trajectory Calculus (QTC) to model human-robot spatial interaction HRSI [55, 59, 60, 61]. QTC use points as primitives in order to represent both the human and the robot, and their relative motion is expressed in a set of tuples of qualitative relationships.

[62] proposed qualitative social rules for robots to have a polite pedestrian behavior while navigating. They used the relative orientation calculus $OPRA_4$ to formalize polite navigation rules in situations such as: crossing, bottleneck or narrow passages, passing groups on the outside or crossing them if they are too large, etc. They simulated motion planning and pedestrian behavior using JWalkerS and SparQ toolbox[1] to investigate how traveling time is influenced by being polite (i.e. following social norms, etc.). However, they did not deal with spatial arrangements of a robot interacting with a group of people (i.e. carrying a joint action). Then, the same authors [63] modeled these pedestrian rules in QLTL (Linear Temporal Logic with Qualitative Spatial Primitives) and presented one exemplary case study using a Kinect camera and a laser range scanner on a mobile robot.

6.2 A Qualitative Spatial Approach to Describe Group-Robot Interactions (QS-GRI)

In this section, an approach to represent the qualitative spatial arrangements for group-robot interactions defined by [37] is presented. First, an iconic representation is given (Section 6.2.1); next, using this representation, the F-formations are described: vis-a-

[1] SparQ toolbox: `http://www.sfbtr8.uni-bremen.de/project/r3/sparq/`

vis (Section 6.2.2), L-shape (Section 6.2.3), circular (Section 6.2.4), horse-shoe (Section 6.2.5), side-by-side (Section 6.2.6), performer-audience (Section 6.2.7).

6.2.1 Iconic Representation

From the point of view of spatial features, interactions between robots and people depend on two factors: (i) distance and (ii) orientation. People are oriented entities in space, which *front* is indicated by their eyes. So, robots need to know that social conventions indicate that to talk properly to somebody they must try to make eye contact with them, which is more feasible if robots approach people from the *front*.

Moreover, robots must be aware that people's personal space usually is not interfered by other people unless they are family, and this space is not allowed to be interfered by robots. So, an interactive distance for a robot is that which is not too close to any person but not too far away for them. [37] defined the *o-space* as the space where people can interact and manipulate shared objects. Similarly, in psychology, peripersonal space is defined as the space wherein individuals manipulate objects, whereas extrapersonal space –which extends beyond the peripersonal space– is defined as the portion of space relevant for locomotion and orienting [64, 65]. Therefore, two individuals that share their peripersonal space can be considered to have an interaction.

In this section, the iconic representation of an individual (robot or person) is shown in Figure 6.1.

Figure 6.1: Iconic representation of a person showing its personal space (ps) and peripersonal space (pps).

That is, any individual fills an area in space (in blue), and (s)he has a personal space (in red) which is private, and a peripersonal space (in green) which is that space that (s)he can reach using their body or a tool. The rest of the white space is the extrapersonal space.

Any person distinguishes spatial locations inside his/her personal and peripersonal space. These areas are usually named as: *front, back, right* and *left*. A person is also an oriented entity in space, defined by his/her *front* where his/her eyes are located. The width of the personal space (ps) depends on the person, their social abilities and culture. Some people would need a wider personal space than other people. These areas can be customized according to the individual person but also parameterized based on psychological experimental studies [66]. The peripersonal space (pps) is dynamic and adaptable, depending on the tool used by the person/robot and their abilities (i.e. flexibility of legs/arms for a person, actuator possibilities in a robot, etc.)

6.2.2 Vis-a-vis Formation

In the vis-a-vis formation by [37], individuals are facing each other. A spatial situation suitable for interaction is defined as that situation in which individuals share part of their peripersonal space. In the vis-a-vis formation, the peripersonal spaces intersects in the front area of both individuals, as it is shown in Figure 6.2. Note that the front of each individual must be turned about 180° to be transformed into the other individual perspective.

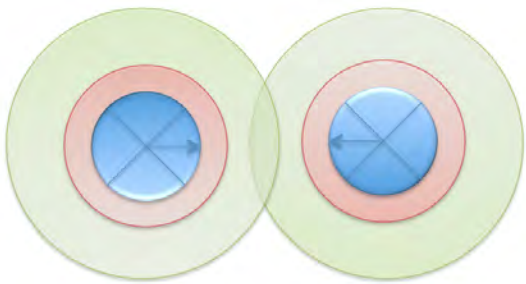


Figure 6.2: Vis-a-vis formation: 180° relationship.

6.2.3 L-shape Formation

In the L-shape formation by [37], two individuals are facing an object having 90° or L-shape separation between them (see Figure 6.3). These two individuals must share some peripersonal space between them. The intersection of this peripersonal space intersects at their *front-left* area of one individual and at the *front-right* area of the other individual. The object observed is not animated, so it has not personal or peripersonal space. The object must be located in the front area of both individuals, which is shared.

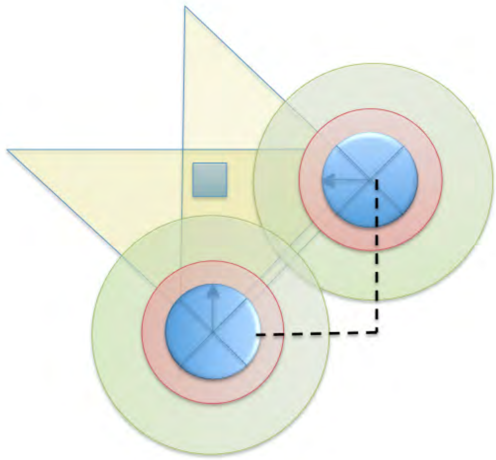


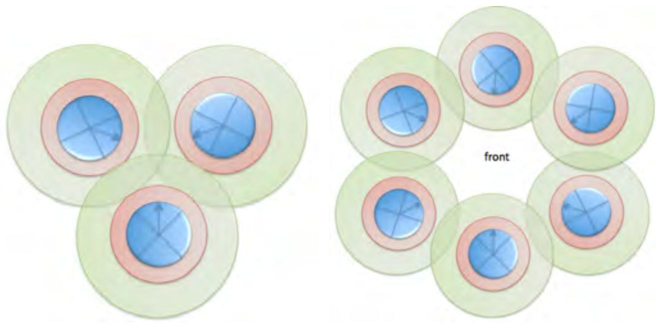
Figure 6.3: L-shape formation.

The individuals are observers, they are not physically interacting with each other, otherwise they would face each other. They are talking about the object. The roles of speaker and listener can be taken in turns. Note that the front of each individual must be turned 90° to be transformed into the other individual perspective.

6.2.4 Circular Formation

The minimal circular formation by [37] is a triangular spatial formation oriented towards the common shared peripersonal space (see Figure 6.4 (a)). In the general case, individuals share their peripersonal space with their neighbors, in the *right* and *left* area. They all share their *front* area (see Figure 6.4 (b)).

The individuals are not only observers, they can interact with each other. The roles of speaker and listener can be exchanged constantly. Note that, in the minimal circular formation, the *front* of each individual must be turned 120° to be transformed into the other individual perspective. In the general circular formation, the front of each individual must be turned $360°/N$ according to the number of individuals, N, to be transformed into the other individual perspective.



(a) (b)

Figure 6.4: Circular formation: (i) minimal circular formation, (ii) general circular formation.

6.2.5 Horseshoe Formation

In the horse-shoe formation by [37], individuals share their peripersonal space with their neighbors, in the *right* and *left* area. They all share their *front* area. The individuals are all of them observers, they are displaced to listen to somebody or to see some object (see Figure 6.5).

Hence, they hold the role of listeners. This is a passive role which can be changed in the case that there is a speaker (usually located at the shared *front*). Note that, in the horse-shoe formation, the *front* of each one of the N individuals must be turned $180°/N$ to be transformed into the other individual perspective.

6.2.6 Side-by-side Formation

In the side-by-side formation by [37], individuals have the same perspective. They share their peripersonal space with their neighbors at their *left* and at their *right*. In the queuing variation, individuals have also the same perspective, but they share their peripersonal space with their neighbors at their *front* and at their *back* (see Figure 6.6). In both cases, individuals' role is passive. They are listeners-observers. Usually, they do not take the speaker roll unless they are given permission for (i.e. for the queuing variation, since they are the head of the queue). Note that, in both side-by-side and queuing formations, as individuals have the same perspective –they are oriented towards the same direction– they must turn 0° to get the same *front* as their neighbors.

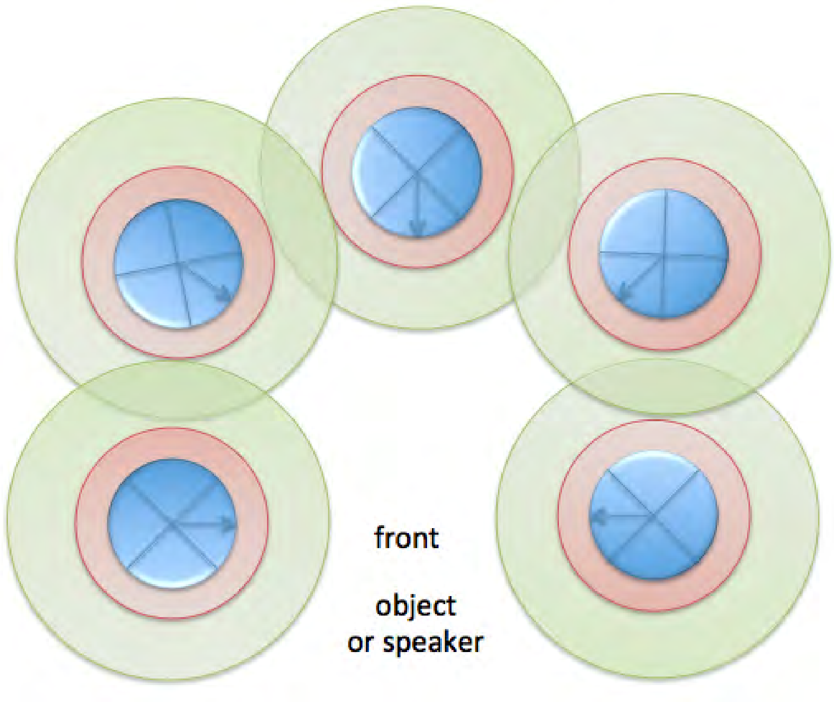
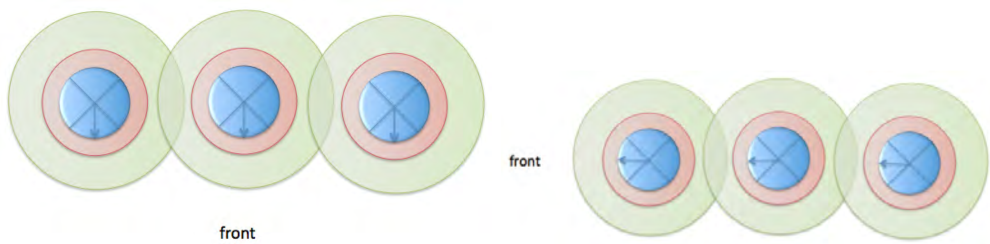


Figure 6.5: Horse-shoe formation.



(a) (b)

Figure 6.6: Side-by-side formation and the queuing variation.

6.2.7 Performer-Audience

All the individuals have the same perspective and they share their peripersonal space with their neighbors at their *front*, *right*, *left* and at their *back* (see Figure 6.7). Their role is passive. They are listeners-observers. They do not take the speaker roll unless they are given permission for, that is, they are asked.

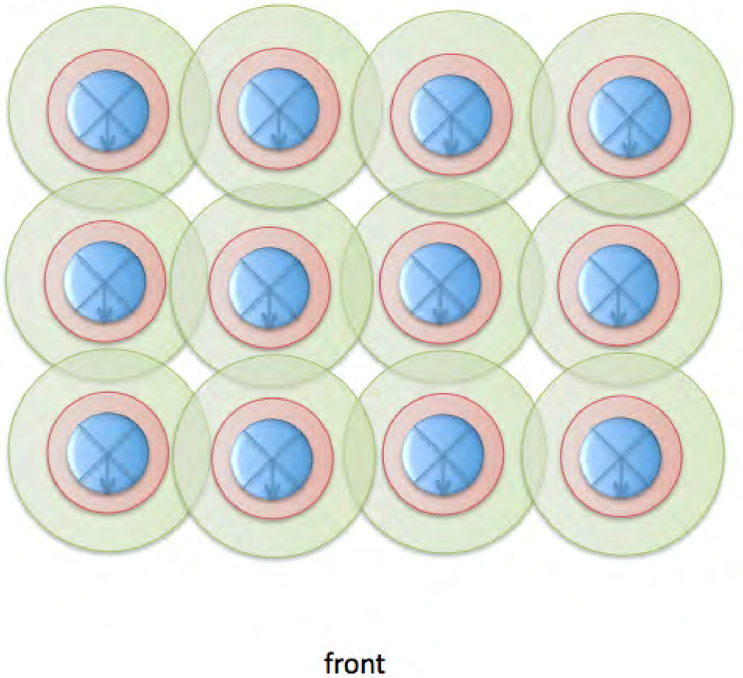


Figure 6.7: Performer-audience formation formation.

6.2.8 Conceptual Neighborhood Situations

In previous sections, we have observed how the Qualitative Spatial model for Group Robot Interaction (QS-GRI) defines the Kendon-formations depending on: (i) the relative location of the individuals involved in the interaction; (ii) the orientation of the individuals (shared front) or not; (iii) their shared peripersonal distance; and (iv) the role of the individuals (observers or interactive).

In this section, we deal with the following challenge: *where the robot should locate itself if its goal is to be included in a group? and towards which direction should it be oriented?*

In order to approach this challenge, the evolution of Kendon-formations between them must be studied. That is, how one formation is transformed into another. These transformations can depend on the role that the robot have, and on the amount of people involved.

Figure 6.8 shows a situation in which the goal of the robot is to interact with one person, and the Kendon-formation selected for the robot to start this interaction: vis-a-vis.

(a) (b)

Figure 6.8: How the QR-GRI is evolving from an initial situation –where an individual is alone– to a vis-a-vis formation.

Figure 6.9 shows a situation in which the goal of the robot is to interact with two people who are placed in a vis-a-vis situation, and which is the Kendon-formation selected for the robot to start this interaction, that is, the minimal circular formation.

(a) (b)

Figure 6.9: How the QR-GRI is evolving from a vis-a-vis situation –where two individuals are interacting– to minimal circular formation.

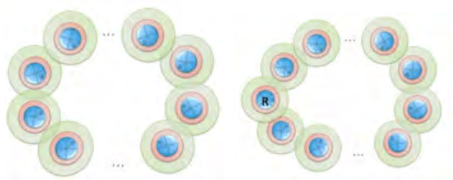
Figures 6.10, 6.11 and 6.12 show situations in which the goal of the robot is to be involved in a group of people who interact among themselves. The initial situation is a group of 3 people situated in a minimal circular formation, and the evolving situations are those where the circle is getting bigger (4-circular formation, 5-circular formation, n-circular-formation).

(a) (b)

Figure 6.10: How the QR-GRI is evolving from a minimal circular situation –where 3 individuals are interacting– to a 4-circular formation.

(a) (b)

Figure 6.11: How the QR-GRI is evolving from a 4-circular situation –where 4 individuals are interacting– to a 5-circular formation.

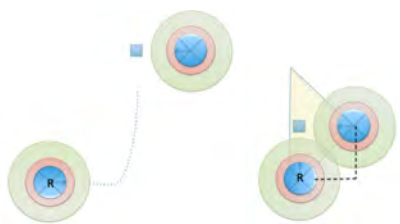


(a) (b)

Figure 6.12: How the QR-GRI is evolving from a n-circular situation –where individuals are interacting– to another n-circular formation.

In situations in which individuals are not interacting with each other, some of them take the role of observers or listeners. In this situations, the following Kendon-formations are suitable for the robot to place itself.

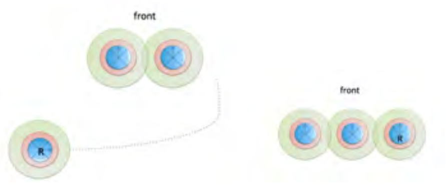
Figure 6.13 shows a situation in which the goal of the robot is to interact with one person while observing an object, and the Kendon-formation selected for the robot to start this interaction: L-shape.



(a) (b)

Figure 6.13: How the QR-GRI is evolving from an initial situation –where an individual is observing an object– to a L-shape formation.

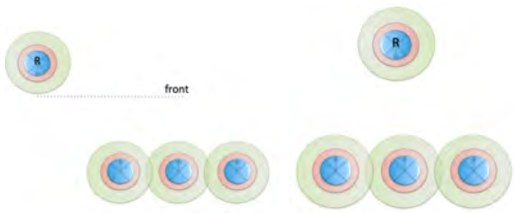
Figure 6.14 shows a situation in which the goal of the robot is to be involved in a group of people which are observing something or someone with whom they cannot interact (i.e. in a performance). These two people are located in a side-by-side formation, and the robot incorporates itself in this side-by-side formation.



(a) (b)

Figure 6.14: How the QR-GRI is evolving from a 2-side-by-side situation –where two individuals are observing someone or something sharing its left/right peripersonal space– to a 3-side-by-side formation which includes the robot.

Figure 6.15 shows a situation in which the goal of the robot is to perform some speech to a group of people who are located in a side-by-side formation. The robot chooses to locate itself at the front.

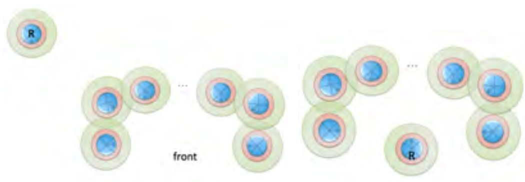


(a) (b)

Figure 6.15: How the QR-GRI is evolving from a 3-side-by-side situation –where 3 individuals are observing someone or something sharing its left/right peripersonal space– to a 3-side-by-side formation which includes the robot at the front.

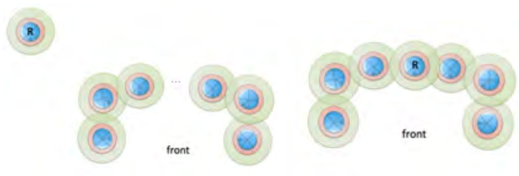
Figure 6.16 shows a situation in which the goal of the robot is to perform some speech to a group of people who are located in a horse-shoe formation. The robot chooses to locate itself at the front. While in Figure 6.17, the goal of the robot is to hear some speech by somebody else or to observe something, then the robot chooses to locate itself among the people. The robot shares its left and right peripersonal space with its neighbors.

Figure 6.18 shows a situation in which the goal of the robot is to perform some speech to a group of people who are located in a performer-audience formation. The robot chooses to locate itself at the front. While in Figure 6.19, the goal of the robot is



(a) (b)

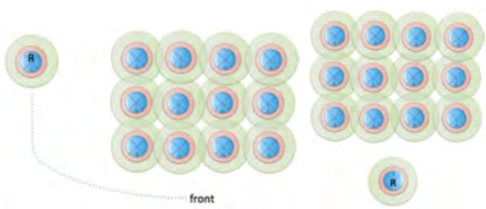
Figure 6.16: How the QR-GRI is evolving from a horse-shoe situation –where individuals are observing someone or something sharing its left/right peripersonal space and also its front– to a horse-shoe formation which includes the robot at the front.



(a) (b)

Figure 6.17: How the QR-GRI is evolving from a horse-shoe situation –where individuals are observing someone or something sharing its left/right peripersonal space and also its front– to a horse-shoe formation which includes the robot among the individuals.

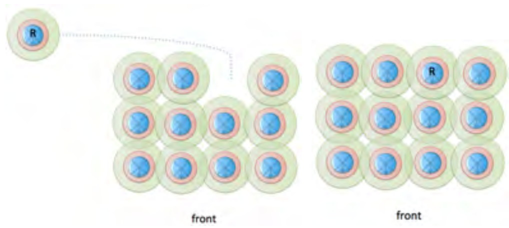
to hear some speech by somebody else or to observe something, then the robot chooses to locate itself among the people. In this case, the robot can have 2 left-right-neighbours and up to 4 neighbours. In the situation depicted, the robot must also share its front peripersonal space is shared with the person in front of it.



(a) (b)

Figure 6.18: How the QR-GRI is evolving from a performer-audience formation –where individuals are observing someone or something– to a performer-audience formation which includes the robot at the front.

All these Kendon-formation transformations have been summarized in Table 6.1. Note that a change of the robot activity/role involves a change in its location in the



(a) **(b)**

Figure 6.19: How the QR-GRI is evolving from a performer-audience formation –where individuals are observing someone or something– to a performer-audience formation which includes the robot among the individuals.

corresponding formation (see lines in Table 6.1), while adding a new person in the group also make the formation to evolve to a different one (change in columns in Table 6.1).

Table 6.1: Table of conceptual neighborhood situations.

	Performer	**Observer**	**Interactive**
1 person	vis-a-vis	L-shape	vis-a-vis
2 people	at front in: side-by-side or minimal circular	L-shape	minimal circular
3 people	at front in: side-by-side or horse-shoe	observer in: side-by-side	circular
4 people	at front in: side-by-side or horse-shoe	observer in: side-by-side horse-shoe	circular
5 people	at front in: side-by-side or horse-shoe	observer in: side-by-side horse-shoe	circular
N people	at front in: side-by-side, horse-shoe or performance	observer in: side-by-side horse-shoe or performance	circular

6.3 Discussion

There are several studies analyzing the spatial interactions from a quantitative approach, expressing spatial relationships in terms of distances and absolute orientations [2]. Since the distances and directions are constantly changing, the representation of the interaction based on these primitives are complex.

The use of Qualitative Spatial Representation techniques can help to abstract and model HRSI. A first approach to use Qualitative Trajectory Calculus HRSI represent qualitatively, using points as primitives to identify the person and the robot in a one-by-one interaction type. Since in real situations HRI can present a great variability in the size of the group, it is possible to use this technique for these cases?

[67] and [62] divided the robot space following proxemics, and they divided the space in: intimate, personal, social and public. In this work, we propose a more psychological point of view dividing the space in personal and peripersonal, which is more related to Kendon definition of o-space [37], where people can interact and manipulate shared objects. Our representation is envisioned to be applied in future human-robot collaboration (HRC) scenarios [68].

As far as we are concerned, there is no previous works in the literature that study the change/evolution of Kendon-formations to help robots to locate themselves following a social convention depending on the role they are assigned (performer/guide, observer/listener, or interactive).

6.4 Conclusions and Future Work

QSR techniques can be a valuable tool for modeling and representing human-robot spatial interactions. In previous works by the authors, a brief analysis was carried out on types of interactions proposing the use of points/regions as primitives to represent the people and the robot [54].

In this chapter, a Qualitative Spatial model for Group Robot Interaction (QS-GRI) is presented which defines Kendon-formations depending on: (i) the relative location of the robot with respect to other individuals involved in the interaction; (ii) the orientation of the individuals (shared front) or not; (iii) the shared peripersonal distance; and (iv) the role of the individuals (observer, performer or interactive).

The evolution of Kendon-formations between them must be studied. That is, how one formation is transformed into another. These transformations can depend on the role that the robot has, and how many people were involved in the interaction, context and space .

As future work we intend to validate this definition to different types of group-robot interaction in real environments. We can use the data from an exploratory study of HRI as a guide robot exhibition in a cultural center.

Chapter 7

Conclusions and Future Work

In this chapter the main conclusions of this research and development around the human-robot interaction in naturalistic environments are shared with the scientific community. Likewise, future works in both research and technological development are presented at the end.

7.1 Conclusions

With the main objective of studying human-robot interactions in naturalistic environments, several research and development were carried out in this thesis, which can be grouped into 4 stages.

Initially, three HRI studies with the REEM robot were carried out in the Cosmo-Caixa science museum: two with the REEM robot as a museum guide and a third study with the robot as an assistant teacher.

In the studies of HRI with the REEM robot as a museum guide, the study was to characterize the behavior of the human-robot groups based on certain characteristics, such as the composition, size and spatial arrangement of the group, observed during the guiding process. Systematic observation technique were used to describe the HR group interaction. During two HRI experiences, REEM robot succeeded in developing the role of a museum guide. Additionally, in the first study registered face-to-face interactions show untrained visitors social behavior addressed to the robot including eye contact, smiles, and greetings.

In both studies, a great percentage of larger HR groups were observed , while the lowest percentage were dyads. This can be explained by the profile of people who visits the science museum that are mostly accompanied.

Likewise, the observed spatial arrangements corresponded to the sizes of the groups, with the performer-audience and leader-follower arrangements being the ones that were found more frequently in larger groups, both in face-to-face and walking group interactions, respectively.

Throughout the guide process we could observe a great dynamic in HR groups, both in size and spatial arrangement. In most cases, the groups behaved in a manner consistent with the established routes. However, certain physical characteristics of the environment as well as the degree of occupancy of the visitors in certain areas tends to modify the HR group behavior.

There are situations where it was difficult to describe a group, either by their composition, size or in their spatial formation, as in crowded environments or in the transitions from a face-to-face group interaction to a walking group behavior and vice versa.

It should be noted that during several transitions from performer-audience to walking group behavior, we observed a special type of arrangement, the "making a corridor", where the robot breaks the O-space and the individuals must make a space for it to pass. In the design of the HRI the robot should take the necessary actions so that individuals can correctly interpret this behavior.

Spontaneous spatial arrangements during guidance may not be effective when confronted with robot's affordances and navigation constraints and therefore limit the system performance. However, social robot behavior can be improved by the robot through new forms of verbal or nonverbal communication.

In a third of HRI with the REEM robot as a teacher assistant, the social presence and identity of the robot was evaluated. In general, the assessment of the social presence of the REEM robot was above the average, being slightly higher the assessments made by female participants. We believe the perceptions of the social presence in this experience could have been stronger if within the activities carried out by the children, the role of the robot would demonstrate a more helping service towards the participants.

Regarding the identity of the robot assigned by the children, the names that the children assigned to the robot mostly did not correspond to human names. From this, a higher percentage of names were mechanical compared to the names of pets. The age assigned by the participants mostly corresponded to the age of a child or an adult. A more in-depth study could try to investigate whether the age indicated by the child corresponded to a desired age of the robot or to an age based on the size of the robot. Despite the fact that the REEM robot has a rather neutral appearance, his voice apparently played a large part of the decisions to indicate that REEM was female. In the case of female participants, this percentage increased. All these issues gives us a clue about the importance of assigning a voice to the robot in the HRI design. To enrich the knowledge of the social presence of a robot, in the future it would be interesting to be able to study this on more diverse age groups.

In order to carry out more studies of HRI in naturalistic environments, a second stage describes the development of an experimental robotic platform for the study of human-robot social interaction, called MASHI. For its development the priority was given to the robot social service, which established objectives of construction and programming of the robotic platform. Therefore, different cycles of HRI study and development of the robotic platform were carried out. Initially, the initial robotic platform (MASHI v00) is described both in terms of hardware and software.

In the first version of the robot (MASHI V01) there are improvements in the structure of the robot. Lowering the weight of laptop to the base, as well as replacing the square structure with a tube as continuation of the mast, allowed the moment of inertia to be considerably reduced. These modifications, coupled with a smoother transition between slogans of movement of the base through programming, allowed the robot to be safer than in the previous version. Although the robot remains still with a mechanical appearance, anthropomorphic proportions of the head were considered in the position of both, the display and the speaker. To validate this very basic robot two brief studies of human-robot interaction have been realized. The exploratory studies were conducted to analyze the types of individuals and groups that interacted with MASHI. In addition, as a result of these interaction experiences, important lessons have been learned about the design of the MASHI robotic platform. There are major challenges in the mechanical and behavioral design so that MASHI increases acceptance and adaptation in social environments.

The second version of the robot (MASHI v02) corresponds mainly to the implementation of the head of the robot as well as the programming of its movements, and this implied improvements for the robot appearance and its non-verbal communication.

However, due to safety issues of the mechanical structure of the head, movements of the head are very limited in the range of rotation and the speed of the servomotors.

The MASHI v02 presented at the Smart City Expo World Congress 2014 caught the attention of many visitors. However, there were technical problems related with the pitch robot head motion that in a certain moment was blocked. Another problem related to the mechanical structure of the head was that several parts of the head joints were very forced.

MASHIv03 presents several improvements both mechanical and in programming areas. From the mechanical point of view, the mechanisms of movement and articulation of the head were improved. It also improved the appearance of the head in general, giving a more stylized appearance. For better control of both the movement of the base and the head, data frame formats were implemented. In this version it was also possible to incorporate a second camera. Also in this version, an artificial comic face, called RobotIcon, have been developed. One goal was to give MASHI a face of its own, with which the robot can express emotions. These emotions can be controlled by a human operator through a browser interface. It is possible to display six different emotions and adjust the corresponding intensity using the face. Later, a survey was conducted to evaluate the recognition value of the facial expressions as well as to evaluate whether displaying the full face or rather just the eye-region affects this recognition rate. The evaluations has shown, that it is highly beneficial to display the whole face in order to transmit all emotions properly, especially for showing happiness.

Finally, the development of MASHI v04 was presented, which consisted mainly in the construction and programming of the robot arms. Wrapped in a process of continuous improvement based on the service provided by the robot, different evolutions were made in the design of the robot arms.

A third stage consists of several HRI studies with the MASHI robot carried out in a cultural center and in an international fair. Specifically, 3 HRI studies carried out at the La Bòbila Cultural Center and 2 at the Smart City Expo World Congress - SCEWC were described.

At La Bòbila Cultural Center, the first study on group-robot interaction was carried out in order to observe visitor's preference, their characteristics and their behaviors. The robot succeeded in developing roles as an exhibition guide (the exhibition was called "(2)"), playing with people and maintaining dialogues. 32 interactions were observed and analyzed. The analysis was focused on visitors as groups more than as individuals. Groups were described according to their age and size, while the behavior were analyzed in terms of f-formations and proxemic behavior. Observational methods applied to evaluate group-robot interaction provide fruitful insight to understand the group-robot interaction by means of human-robot spatial relationships. On the other hand, the survey conducted, although moderate in number, yielded interesting results that will help us in future designs of the robot and the HRI in general.

A second study of HRI at La Bobila with MASHI as an exhibition guide (exhibition “Vaixells a la mar”) was carried out to evaluate the perceptions that visitors have towards the robot. In this study results obtained in the anthropomorphism section qualifies MASHI robot as natural, having consciousness, alive and moving elegantly. In spite of this, there is a greater percentage of those who consider that the robot is more similar to a machine than a human. This appreciation corresponds to the fact that what we want to enhance is the service offered by MASHI to the user, beyond of the aesthetic appearance of it. The animacy section shows that a greater percentage of the participants think that the robot is more organic, more lifelike and more interactive. It should be noted the there exists a large percentage of participants who see the robot as lively. Talking about likeability, this feature got a good rating. The vast majority of participants liked the robot, while an important majority thinks that the robot is friendly, kind, pleasant and nice. In perceived intelligence of the robot also obtained a very good valuation in all its items; mainly in intelligence and be responsible, followed by knowledgeable, sensitive and competent.

A third HRI study carried out in La Bòbila was reported. In this we studied both the spatial formations of the groups during the interaction as well as the perceptions of the people who interacted with the robot. From the spatial arrangement analysis performed, it is quite hard to find a robust and evidence-based reply to the research questions posed. MASHI is absolutely able to draw people’s attention throughout different age groups. Nonetheless, it is not able to transfer this attention to the visit and it is not capable of keeping people engaged for a long time. MASHI elicits two different types of social HRI depending on the age group of its interactants: a more courteous and relaxed one from adults, and a very intrusive and active one from children. Such a differentiation is made evident by proxemics. Indeed, children invade MASHI’s intimate space, whereas adults always maintain themselves at a personal distance. The navigation of the robot is safe, but the intentions of the robot are difficult to argue for the crowd. To solve such an issue, the robot could either use verbalization or environment affordances (following the succession of glass cases) to make displacement smoother.

The role of the robot during the guide is not clear, it should take a stronger stance towards the crowd, making its role, leading and guiding, and inner states transparent. In this last case for example, the robot could use its facial expression or a red screen color to express its frustration when its hands are shaken in an improper way.

Conversation could be used to keep the attention of audience lively. Creating a character around a robot and presenting its story could be either a way to make children more interested to what the robot is saying. What is more, to enhance the quality of the interaction, the robot must use its body. Hand-shaking is one of the affordances that could be used during interaction.

Last, we think that F-formations should be used to place a robot within a group in a social manner. Theories more related to crowd or group behaviour and displacement could produce more significant results in crowding contexts.

To sum up, MASHI's configuration is a very good starting point to reach a high quality social HRI, especially with children. Nonetheless, the richness of the interaction the robot is delivering should be improved and other criteria should be used to grasp the real potential of the robot in guiding and engaging people.

Two HRI studies in the Smart City Expo World Congress - SCEWC were carried out with the MASHI robot with the "selfie robot" service. It was really surprising how MASHI got people's attention. Although it was not the best robot of the Congress, nor the most modern, people from all ages were willing to take a picture with MASHI. Some people were really interested and wanted to know more about the project and the robot. However, there were some people that didn't understand exactly the purpose of the robot and the service it was providing. Sometimes people were staying in front of the robot, and someone had to explain to them the task that it was performing, and how to interact with. This drawback gives an idea that's worth taking it into account in future HRI designs.

Another issue that was detected is that most of the times it is very difficult to establish a conversation between the operator and the visitor standing in front of the robot. The operator had to be really fast with the text-to-speech software to catch people's attention and to chat. A possible solution is that the system allows the operator to add dialogue phrases in a more dynamic way. Given the characteristic of the service provided by the robot, in the evaluation of the robot's performance, metrics from social networks were used.

Finally a novel approach to represent the spatial relations of the HR groups using theories from psychology area was proposed. In this stage, a Qualitative Spatial model for Group Robot Interaction (QS-GRI) is presented which defines face-to-face formations depending on: (i) the relative location of the robot with respect to other individuals involved in the interaction; (ii) the orientation of the individuals (shared front) or not; (iii) the shared peripersonal distance; and (iv) the role of the individuals. The evolution of F-formations between them must be studied. That is, how one arrangement is transformed into another. These transformations can depend on the role that the robot have, and how many people were involved in the interaction.

7.2 Future Work

Some possible future works can be generated in the wake of this thesis work, and they are listed below:

- As a teleoperated robotic platform, as the control of the robot becomes more complex, for example with the control of the joints of the robot, it is necessary to provide more support to the operator so that he can handle these complexities. An alternative is the use of alternative HCI such as motion sensing input device to detect the operator's gestures of the arms to control the movement of the robot, or computer vision to detect position and orientation of the head or facial expressions to be replicated in the robot.

  In addition, the control information generated by the operator can be processed through artificial intelligence techniques to reproduce more autonomous social behaviors, freeing the operator of cognitive load.

- In the robotic platform it is also necessary to be able to parameterize certain features, such as the configuration of the operators (*e.g.* the operator gender can be specified a robot gender), robot configuration (*e.g.* the language can be configured depending on the language used by the users that interact with the robot).

- Take advantage of the characteristics of the architecture used in the robotic platform so that those processes where a greater intensity of processing or quantity of information are required can be executed in the cloud.

- To evaluate the social interaction HR in the study, the recordings of video cameras located outside the robot were used. In very large spaces this can generate several difficulties. A possible improvement is to use a camera located in the robot that can capture the entire scene around (*e.g.* an omnidirectional camera). From these images computer vision techniques can be applied for the detection of people, groups of people and their spatial relationships.

- Based on different HRI experiences in different environments and with different types of robots, a QS-GRI database can be created. Applying artificial intelligence techniques robots can reproduce social behaviors automatically.

- Study how to incorporate QS-GRI into the behavior of the robot, either autonomous or teleoperated, so that it can use these formations to position the robot; besides studying for example the reactions and acceptance of people.

Appendix A

REEM Experiences

Cuestionario

Edad: ________

Sexo: ☐**Niña** ☐**Niño**

DE LAS SIGUIENTES FRASES ENCIERRA CON UN CÍRCULO LA RESPUESTA QUE MÁS SE AJUSTE A TU EXPERIENCIA EN ESTA ACTIVIDAD

☹ Corresponde a **total desacuerdo** y ☺ Corresponde a **total acuerdo**

---

1. Al relacionarme con REEM, sentía como si lo hiciera con una persona de verdad.

2. A veces me sentía como si REEM realmente me estuviera mirando.

3. Puedo imaginar al robot (REEM) como un ser vivo.

4. A veces parecía como si el robot tuviera sentimientos

Pon un nombre al robot REEM ______________________

Qué edad crees que puede tener______________________

Crees que es un chico o una chica ___________________

Muchas gracias por tu participación! ☺

Figure A.1: Questionnaire.

Appendix B

MASHI Experiences

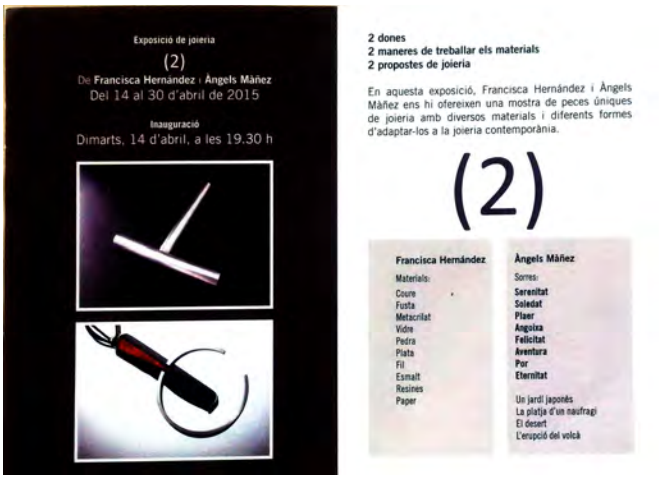
B.1 *"(2)"* Jewelery Exhibition

Table B.1: Jewelery exhibition (2) script

Saludo de bienvenida al Centro Cultural La Bobila
Muy buenas noches, bienvenidos al Centro Cultural La Bobila de LHospitalet de Llobregat.
Soy Dennys, el operador de la plataforma robòtica Mashi, y estamos aquí para presentarles la exposición de joyeria denominada (2).
Por qué 2? Dos mujeres, 2 maneras de trabajar los materiales, dos propuestas de joyeria.
En esta exposición, Francisca Hernandez y Angels Máñez nos ofrecen una muestra unica de piezas de joyeria con diversos materiales y diferentes formes de adaptarlos a la joyeria contemporanea.
Parte 1: arenas (Àngels)
En esta parte de la exposición, denominada arenas, Angels ha querido centrar la exposición en 4 ámbitos definidos por 4 arenas que dibujan una buena parte de las inquietudes y deseos de la vida.
Las piezas de esta exposición son todas hechas a mano, básicamente trabajando la plata y los materiales adecuados a cada uno de los conceptos.
(1) Un jardín zen
Serenidad, espiritualidad, austeridad, equilibrio ...
• Materiales: plata
(2) La playa de un naufragio
Miedo, soledad, angustia, supervivencia, esperanza ...
• materiales encontrados: plata, huesos, coral muerto, cuerno
(3) el desierto
Silencio, inmensidad, luz, vida ...
• Materiales: plata y terracota
(4) Las cenizas de un fuego
Símbolos ancestrales y eternos que emergen del inicio de la vida. Simbología divina o diabólica.
• Materiales: Plata y resina teñida.

Parte 2: Materiales (Paca)
Los materiales de uso común, aplicados a la joyería contemporanea, es el argumento de la propuesta de Paca.
(1)La madera: naturalez, ductibilidad, calidez.
(2)La plata y el cobre: dos elemntos de la tabla periódica, dos temperaturas de fusión distintas que permiten su trabajo como si se tratara de un solo material.
(3)El vidrio: Duro, fràgil, delicado, transparente, de usos cotiadianos y muy diversos,desde un envase en un super a una vidriera en una catedral.
(4)El algodón: Vegetal, cálido, noble, a menudo una segunda piel.

(a)



(b)

Figure B.1: Invitation to the jewelery exhibition (2)

B.2 *"Vaixells a la mar"* Exhibition

B.3 *"Segon Esdeveniment d'Integració Multicultural"* Exhibition

B.4 Smart City Expo World Congress

Cuestionario

Tú eres: ☐Chico ☐Chica **Tu edad?:** ________

Encierra con un círculo la respuesta que más se ajuste a tu experiencia en esta actividad (☺ es total acuerdo y ☹ es total desacuerdo)

1.	Mashi es un buen guía de exhibición en La Bòbila.
2.	Me gusta acompañar al robot durante la exposición.
3.	Me gusta que el robot me siga.
4.	Me gusta seguir al robot.
5.	Me gustaría que Mashi fuera más rápido.
6.	Los movimientos de la cabeza son adecuados.
7.	Me gustan las actividades extras que hace el robot.
8.	Me gusta que el robot reconozca mi nombre.
9.	Me gusta que el robot ponga música.
10.	Me gusta bailar con el robot.
11.	Me gusta abrazar al robot.
12.	No le tengo miedo a Mashi, se ve inofensivo.
13.	Me gusta acercarme mucho al robot
14.	El robot se ve muy pesado.
15.	El robot es muy alto
16.	Me gusta la cara del robot.
17.	Me gustan las expresiones faciales que hace el robot (por ej. felicidad, tristeza, etc.)
18.	Me gustaría que el robot tuviera brazos

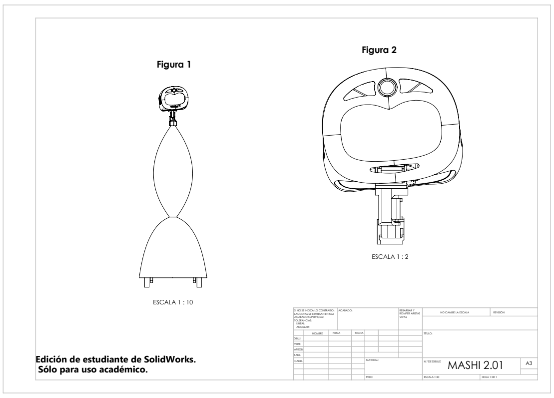
Si Mashi tuviera brazos, ¿para qué crees que podrían servir?________________

____

- Píntale a Mashi! Colorea la Figura 1 de la forma que más te guste!
- ¿Te gustaría que Mashi tenga brazos? Dibújale unos en la Figura 1!
- ¿Te gustaría que Mashi tenga otro rostro, quizás el tuyo? Dibújale y píntale un rostro a Mashi (Figura 2)

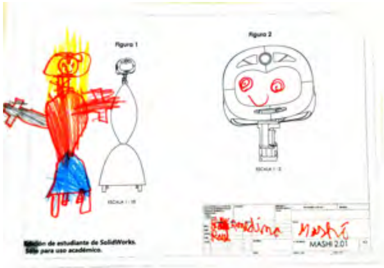
¡Muchas gracias por tu participación! ☺

(a)

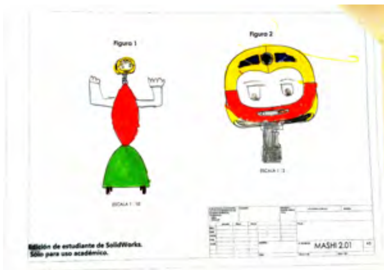


(b)

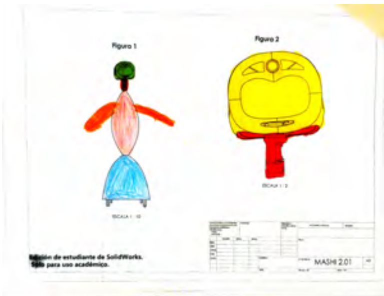
Figure B.2: Jewelery exhibition '(2)' questionnaire



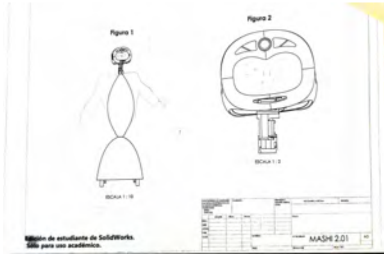
(a)



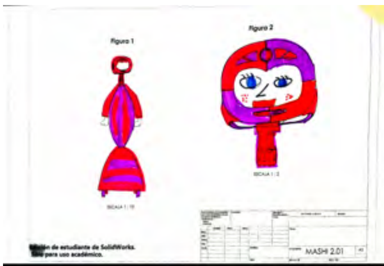
(b)



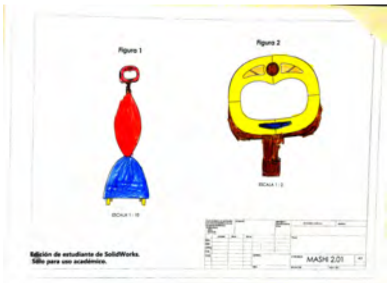
(c)



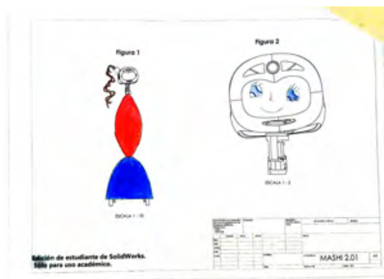
(d)



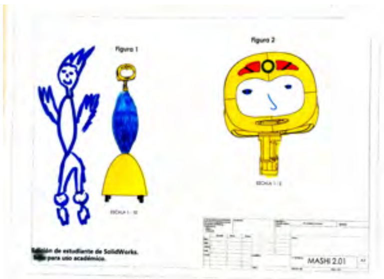
(e)



(f)



(g)



(h)

Figure B.3: Several Jewelery exhibition (2) result draws

(a)

(b)

Figure B.4: Invitation to the exhibition Vaixells a la mar

1. buscar TARGET
2. Si <TARGET ENCONTRADO> entonces
3. Hola, muy buenas tardes. Me llamo Mashi y soy un robot guía de exposición.
4. ¿Desean hacer un recorrido por una interesante exposición?
5. Si <CASO NEGATIVO> entonces
6. (De acuerdo) (muchas gracias!)
7. Ir a 1.
8. Si <CASO AFIRMATIVO>
9. ¡Genial!
10. Si <NO está en POSICION INICIAL> entonces
11. (Acompáñenme por favor!)
12. [ir a POSICION INICIAL]
13. Ir a 15.
14. Si <SI está en POSICION INICIAL> entonces
15. INTRODUCCIÓN: Esta exposición se denomina BARCOS A LA MAR realizada por Tucho. Aunque los barcos aparecen como tema principal de la exposición no es ésta una muestra de modelismo naval. El autor no copia un modelo, sino crea una ficción. Aunque el material de base es la madera no son obras de ebanistería. El material es de desecho de las carpinterías, comprado en sacos para leña. Aunque representan barcos, ninguno navegaría. Queda solo una idea de proas, anclas, velas, hélices, chimeneas, elementos que se combinan de una y mil maneras.
16. (¿Continuamos?)
17. Peana 1: MONTEVIDEO. BENEDETTI. Esta primera peana va dedicada a las raíces del autor: Montevideo, su mar y sus gentes. Tucho eligió la poesía de Mario Benedetti como expresión de ese sentimiento. Un poema está dedicado a la ciudad, Montevideo. Los otros dos, de tema marinero, sirven de inspiración a los barcos que las acompañan e ilustran.
18. (¿Continuamos?)
19. Peana 2: BARCO DE PESCA. La pesca tiene un papel fundamental para el marinero. Por eso el autor representa el mar oscuro, potente y los peces, objetivo del pescador aparecen, en fuerte contraste, con colores vivos y brillantes. El barco grande muestra claramente los elementos que lo componen: una sucesión de maderas verticales, la caseta del pescador dando unidad al conjunto y la chimenea metálica como remate final.
20. (¿Continuamos?)
21. Peana 3: BARCOS DE AUXILIO. Al barco mayor lo define y caracteriza la pieza con el hueco del ancla (anzuelo) y la curva suave que define la proa. Tucho la mantiene tal cual, como la encontró entre los retales. El resto son simples tacos de madera prismáticos. La hélice propulsora es el ventilador de un ordenador ☺

22. (¿Continuamos?)
23. Peana 4: BARCOS A VAPOR. El barco superior está definido por un gran plano vertical animado por la red que dibujan los remaches del casco y las escotillas de la nave. La caseta es el elemento diferencial, sobresaliente, que juega con el resto del conjunto.
24. (¿Continuamos?)
25. Peana 5: BARCOS DE CARGA. En esta peana cuenta mucho el color que dan los corchos. Representan el motivo principal del transporte marítimo y de ahí su relevancia en la representación.
26. (¿Continuamos?)
27. Peana 6: BARCOS ANTIGUOS. Ejemplo de una misma escuadría de madera en la sucesión y remate superior, solo diferenciadas por tonos de barniz. La quilla es la que unifica esta sucesión. Las velas potencian el sentido vertical de la composición.
28. (¿Continuamos?)
29. Peana 7: BARCOS A VELA. Mientras en los pequeños las velas son livianas y coloridas, en el mayor la vela es tan maciza como el propio casco. Contraste de expresión para la representación de un mismo elemento.
30. (¿Continuamos?)
31. Peana 8: BARCOS DE PASAJEROS. El transatlántico, el Titanic o la golondrina. El océano o el mar más próximo. Aquí el protagonista es el viajero. Buen viaje! Y buenas tardes!
32. (¡muchas gracias!)
33. ¿Desean conocer más sobre el autor?
34. Si <caso POSITIVO> entonces:
35. Sobre el autor: Nacido en Montevideo (Uruguay). Hace 43 años que Tucho Bergeret vive en Barcelona. De profesión arquitecto, hasta que la crisis lo jubiló. Su mujer también es uruguaya y arquitecta. Ambos llegaron aquí con un Grupo de Viaje de Estudiantes de Arquitectura y se quedaron. Sus hijas, Karol y Sesi, son catalano-uruguayas (con las dos nacionalidades). Son sus promotoras, una en la parte artística y montajes, y la otra en la difusión a través de las redes sociales. Comenzó a trabajar en el tema de los barcos hace dos años, a la vuelta de un viaje a Montevideo. Recuerdos de la playa, la Rambla, el Puerto, sus barcos y sus gentes; y una partida de madera con cortes diagonales en los que vió las proas de barcos. Un saludo de parte de Tucho.
36. Ir a 38.
37. Si <caso NEGATIVO> entonces
38. Una cosa más ¿Me pueden ayudar llenando un breve cuestionario sobre mi desempeño por favor?
39. Si <caso POSITIVO> entonces

40. (¡acompáñenme por favor!)
41. Aquí encontrarán el cuestionario,
42. (muchas gracias)
43. Ir a 1
44. Si <caso NEGATIVO> entonces
45. De acuerdo
46. muchas gracias!

(a) (b) (c)

Figure B.5: Vaixells a la mar script

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Cuestionario: Robot guía de exposición

Sexo: ☐Masculino ☐Femenino Edad: ______

ANTROPOMORFISMO

Por favor, califica tu impresión del robot en estas escalas

Imitación, no genuino, artificial	1	2	3	4	5	Natural
Parece máquina	1	2	3	4	5	Parece humano
No tiene conciencia	1	2	3	4	5	Tiene conciencia
Es artificial	1	2	3	4	5	Tiene vida
Se mueve rígidamente	1	2	3	4	5	Se mueve con elegancia

ANIMACIDAD

Por favor, califica tu impresión del robot en estas escalas

Muerto	1	2	3	4	5	Vivo
Inactivo	1	2	3	4	5	Animado
Mecánico	1	2	3	4	5	Orgánico
Artificial	1	2	3	4	5	Vivo y real
Inerte	1	2	3	4	5	Interactivo
Es indiferente	1	2	3	4	5	Reacciona con entusiasmo

SIMPATÍA

Por favor, califica tu impresión del robot en estas escalas

No me gusta	1	2	3	4	5	Me gusta
Es antipático	1	2	3	4	5	Es amistoso
No es amable	1	2	3	4	5	Es amable
Es desagradable	1	2	3	4	5	Es agradable
Es horrible	1	2	3	4	5	Es bonito

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

INTELIGENCIA PERCIBIDA

Por favor, califica tu impresión del robot en estas escalas

Es incompetente	1	2	3	4	5	Es competente
Es ignorante	1	2	3	4	5	Es experto
Es irresponsable	1	2	3	4	5	Es responsable
No es inteligente	1	2	3	4	5	Es inteligente
Es tonto	1	2	3	4	5	Es acertado

SEGURIDAD PERCIBIDA

Por favor, califica tu impresión del robot en estas escalas

Ansioso	1	2	3	4	5	Relajado
Agitado	1	2	3	4	5	Calmado
Quieto	1	2	3	4	5	Sorprendido

¡Muchas gracias por su colaboración!

(a) (b)

Figure B.6: Vaixells a la mar questionnaire

Figure B.7: Invitation to the exhibition.

Guión

INTRODUCCIÓN:

La Asociación Iberoamericana Solidaria organiza un evento multicultural, el evento cuenta con la presencia de artistas de diferentes países, entre ellos de los países iberoamericanos y como país invitado Bulgaria.

Esta exposición pone en claro el carácter transversal del proceso histórico de la escultura ancestral y milenaria de los Andes, donde se reflejan las culturas del planeta tierra

Peana 1-4: ESCULTURA NATIVOS UNIVERSALES "RUNAS ALLPAMANTA".

AUTOR: AMARUK KAYSHAPANTA

TECNICA: AMAWTIKA-MADERA

AÑO: 2003

P1: América; P2: África; P3: Indonesia; P4: China

En estas sección se presenta una panorámica de esculturas en las que "la madera, sus calidades y sus propiedades plásticas fueran elementos constitutivos", cuyas obras tratan de expresar los lenguajes ancestrales de los escultores originarios de "América"

Así, en el aspecto explicativo, esta exposición pone en claro el carácter transversal del proceso histórico de nuestra escultura vanguardista en madera, en la que vemos cruzarse disciplinas diversas -entre lo escultórico, lo pictórico y lo gráfico- y combinarse aportaciones de lenguajes muy distintos, en la obra del artista multidisciplinar y afamado como director de cine y escritor, Amaruk Kayshapanta

Un primer panorama de la escultura en madera Kayshapanta logra invitarnos a un viaje por el cosmos de este universo y sus orígenes cuya facilidad se debe a sus años en la práctica de la escultura en madera, cuyo horizonte comprobamos extraordinariamente amplia universalidad la del artista indígena de Kichwa de los Andes.

Peana 5-8: Artesanías de Ecuador

La asociación sociocultural General Eloy Alfaro expone vitrinas con artesanías elaboradas en Ecuador.

Peana 5: Ecuador, país megadiverso

(a)

En esta vitrina se representa la gran diversidad cultural y artística del Ecuador, puesta a manifiesto en diferentes tipos de artesanías. En el centro se puede apreciar una escultura en madera que representa la mitad del mundo. El abanico y la botella son artesanías elaboradas en la costa ysierra, en el lado izquierdo unos bolsos y una banda multicolor representan la indumentaria e instrumentos de música andina, tales como la alpargata, la quena, el tambor y la flauta de pan. En el lado derecho un coche multicolor, la chiva, representa la cultura montubia rica en productos tropicales. La tortuga representa la diversidad marina del Ecuador, con Las Islas Galápagos como su mayor exponente.

Peana 6: Artesanías en paja toquilla

El material de estas artesanías, la paja toquilla, procede principalmente de la zona rural del cantón Montecristi, donde se dan las condiciones ideales para su crecimiento. Sus artesanos son muy reconocidos a nivel mundial por sus variedades de artesanías utilitarias y sobre todo por su especial trabajo en la confección del sombrero fino de paja toquilla conocido mundialmente como sombrero panamá.

Peana 7: Cultura Afroamericana

En esta vitrina se pueden apreciar artesanías que representan la cultura afroamericana radicada principalmente en la provincia de Esmeraldas, al norte del Ecuador.

Peana 8: Cultura Indígena

Como parte de un Estado Pluricultural, multirracial y Multiétnico como lo es Ecuador, la cultura indígena remonta su existencia desde hace unos 13000 años, y su extensión sobrepasa las fronteras del país, por ello la presencia de varias artesanías provenientes de Chile, Bolivia, Perú y Uruguay.

Sobre el autor/es:

Son artistas de países de Iberoamérica.

(b)

Figure B.8: Exhibition scripts

Sentencias

Afirmativas: si, así es

Negativas: no, no lo sé, no puedo

Expresiones de duda

Posiblemente

Probablemente

Quizá

Quizás

Tal vez

Interrogativas

¿Continuamos?

¿Por favor, me ayudan llenando un cuestionario que está en la mesa de información?

Respuestas

Me llamo Mashi

Soy un robot

Soy un robot guía de la exposición

Frases coloquiales

- Cortesía

Hola, muy buenas tardes!

Genial!

De acuerdo!

Muchas gracias!

¡Hasta ahora!

- Petición

¡Permiso por favor!

¡Acompáñenme por favor!

!Silencio por favor¡

Alerta - Amarillo

¡Con cuidado!

(a)

¡Con precaución!

Alerta-Rojo

¡Más respeto por favor!

¡Me haces daño!

¡Me tengo que ir!

Expresiones

https://espanol.lingolia.com/es/gramatica/tiempos-subjuntivo/expresiones

http://www.hispanoteca.eu/Foro-preguntas/ARCHIVO-Foro/Subjuntivo%20en%20subordinadas%20sustantivas.htm

(b)

Figure B.9: Exhibition robot dialogue

NAME OF THE OBSERVER: ____________________________

CHECK SHEET FOR SOCIAL HRI EVALUATION

TYPE OF INTERACTOR

	Individual (1 person)
	Couple (2 people)
	Group (3 or more people)

TYPE OF F-FORMATIONS

MINUTE	NONE	CIRCULAR	SEMI-CIRCULAR	L-SHAPE	FACE-TO-FACE	SIDE-BY-SIDE	RECTANGULAR
min 1							
min 2							
min 3							
min 4							
,,,							
min n							

More frequent F-formations: ____

Cohen's Kappa coefficient (inter-rater agreement): ____________________

Figure B.10: Check sheet for analysis.

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Cuestionario: Exposición

(Vitrinas en la entrada principal)

Género: ☐ **Masculino** ☐ **Femenino** **Edad:** ________

Encierra con un círculo la respuesta que más se ajuste a tu experiencia en la exposición (1 es total desacuerdo y 5 es total acuerdo)

1.	He venido al Centro Cultural a ver expresamente la Exposición.	1	2	3	4	5
2.	Tengo el tiempo suficiente para ver todas las obras.	1	2	3	4	5
3.	Es la primera vez que visito la exposición.	1	2	3	4	5
4.	Visito con frecuencia al Centro Cultural.	1	2	3	4	5
5.	Visito con frecuencia la Biblioteca.	1	2	3	4	5
6.	Normalmente vengo solo al Centro Cultural.	1	2	3	4	5
7.	Me ha gustado la Exposición.	1	2	3	4	5
8.	He aprendido de la Exposición.	1	2	3	4	5
9.	He conocido sobre los materiales con que están hechas las obras.	1	2	3	4	5
10.	He conocido sobre la procedencia de las obras	1	2	3	4	5
11.	He conocido sobre las técnicas utilizadas para realizar las obras.	1	2	3	4	5
12.	He conocido sobre la biografía del autor	1	2	3	4	5
13.	He conocido sobre las motivaciones del autor para realizar las obras.	1	2	3	4	5
14.	He conocido de alguna anécdota que haya tenido el autor con alguna de sus obras.	1	2	3	4	5

Por favor, conteste las siguientes preguntas:

- ¿Cuántas veces ha visto anteriormente la Exposición?

1	2	3	4	5

- ¿Cuántas vitrinas ha visitado ahora?

1	2	3	4	5	6	7	Todas

- ¿Cuántos minutos cree que son suficientes para recorrer toda la exposición (ubicada en el hall principal)? ______
- ¿Cuántas personas le han acompañado? ______

¡Muchas gracias por su participación!

Figure B.11: Questionnaire without the robot

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Cuestionario: Robot guía de exposición

Género: ☐Masculino ☐Femenino **Edad:** _________

Encierra con un círculo la respuesta que más se ajuste a tu experiencia en la exposición (1 es total desacuerdo y 5 es total acuerdo)

1.	He venido al Centro Cultural a ver expresamente la Exposición.	1	2	3	4	5
2.	Tengo el tiempo suficiente para ver todas las obras.	1	2	3	4	5
3.	Es la primera vez que visito la exposición.	1	2	3	4	5
4.	Visito con frecuencia al Centro Cultural.	1	2	3	4	5
5.	Visito con frecuencia la Biblioteca.	1	2	3	4	5
6.	Normalmente vengo solo al Centro Cultural.	1	2	3	4	5
7.	Me ha gustado la Exposición.	1	2	3	4	5
8.	He aprendido de la Exposición.	1	2	3	4	5
9.	He conocido sobre los materiales con que están hechas las obras.	1	2	3	4	5
10.	He conocido sobre la procedencia de las obras	1	2	3	4	5
11.	He conocido sobre las técnicas utilizadas para realizar las obras.	1	2	3	4	5
12.	He conocido sobre la biografía del autor	1	2	3	4	5
13.	He conocido sobre las motivaciones del autor para realizar las obras.	1	2	3	4	5
14.	He conocido de alguna anécdota que haya tenido el autor con alguna de sus obras.	1	2	3	4	5

Por favor, conteste las siguientes preguntas:

- ¿Cuántas veces ha visto anteriormente la Exposición?

1	2	3	4	5

- ¿Cuántas vitrinas ha visitado ahora?

1	2	3	4	5	6	7	Todas

- ¿Cuántos minutos cree que son suficientes para recorrer toda la exposición (ubicada en el hall principal)? ______
- ¿Cuántas personas le han acompañado? ______

(Continúa en la siguiente página…)

(a)

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

Cuestionario: Robot guía de exposición (Cont.)

Por favor, para cada apartado valora tu impresión del robot utilizando las escalas numéricas

ANTROPOMORFISMO

Imitación, no genuino, artificial	1	2	3	4	5	Natural
Parece máquina	1	2	3	4	5	Parece humano
No tiene conciencia	1	2	3	4	5	Tiene conciencia
Es artificial	1	2	3	4	5	Tiene vida
Se mueve rígidamente	1	2	3	4	5	Se mueve con elegancia

ANIMACIDAD

Muerto	1	2	3	4	5	Vivo
Inactivo	1	2	3	4	5	Animado
Mecánico	1	2	3	4	5	Orgánico
Artificial	1	2	3	4	5	Vivo y real
Inerte	1	2	3	4	5	Interactivo
Es indiferente	1	2	3	4	5	Reacciona con entusiasmo

SIMPATÍA

No me gusta	1	2	3	4	5	Me gusta
Es antipático	1	2	3	4	5	Es amistoso
No es amable	1	2	3	4	5	Es amable
Es desagradable	1	2	3	4	5	Es agradable
Es horrible	1	2	3	4	5	Es bonito

Centre d'Estudis Tecnològics per a l'Atenció a la Dependència i la Vida Autònoma
UNIVERSITAT POLITÈCNICA DE CATALUNYA

INTELIGENCIA PERCIBIDA

Es incompetente	1	2	3	4	5	Es competente
Es ignorante	1	2	3	4	5	Es experto
Es irresponsable	1	2	3	4	5	Es responsable
No es inteligente	1	2	3	4	5	Es inteligente
Es tonto	1	2	3	4	5	Es acertado

SEGURIDAD PERCIBIDA

Ansioso	1	2	3	4	5	Relajado
Agitado	1	2	3	4	5	Calmado
Quieto	1	2	3	4	5	Sorprendido

¡Muchas gracias por la colaboración!

(b) **(c)**

Figure B.12: Questionnaires with the robot

Start

Hi	Welcome to the city of L'Hospitalet
Good morning	You are in a space full of creativity and innovation. If you read our activities program you will see how we shape our city
Good afternoon	

Development

Do you want to take a selfie with me and view the photo in the twitter account Mashi Robot?	Move down, please!
Great!	3, 2, 1...smile!
Fantastic!	Beautiful!
here we go	Wonderful!
Move to my right, please	Thank you very much!!
Move to my left, please!	Enjoy the day!

Phrase

Yes	I understand
Thanks	I don't understand
Very good!	I'm sorry
Fine thanks!	A moment please!
and you?	excuse me!
My name is Mashi	be careful, please!
Nice to meet you	see you
I'm a telepresence robot	bye bye!
No	The twitter account is mashi robot
I don't know	You can access the selfie photo through the twitter account @ Mashi Robot, thanks!
Maybe	

Figure B.13: SCEWC Vocabulary

Bibliography

[1] Marta Díaz-Boladeras, Amara Andrés, Judit Casacuberta Bagó, and Cecilio Angulo Bahón. **Propuesta metodológica para la evaluación de la interacción persona-robot en diversos escenarios de aplicación**. In *Workshop ROBOT 2011 : robótica experimental : Libro de Actas*, pages 617–621, 2011. 1, 7, 9

[2] Annica Kristoffersson, Kerstin Severinson Eklundh, and Amy Loutfi. **Measuring the Quality of Interaction in Mobile Robotic Telepresence: A Pilots Perspective**. *International Journal of Social Robotics*, pages 1–13, 2012. 1, 15, 16, 17, 19, 20, 47, 189

[3] Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. **How close?: model of proximity control for information-presenting robots**. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, HRI '08, pages 137–144, New York, NY, USA, 2008. ACM. 1

[4] David J. Feil-Seifer and Maja J. Matarić. **Distance-Based Computational Models for Facilitating Robot Interaction with Children**. *Journal of Human-Robot Interaction*, **1**(1):55–77, Jul 2012. 1, 19

[5] Markus Kuderer, Henrik Kretzschmar, and Wolfram Burgard. **Teaching Mobile Robots to Cooperatively Navigate in Populated Environments**. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3138–3143, Tokyo, Japan, Nov 2013. 1

[6] Cindy L. Bethel and Robin R. Murphy. **Affective expression in appearance constrained robots.** In Michael A. Goodrich, Alan C. Schultz, and David J. Bruemmer, editors, *HRI*, pages 327–328. ACM, 2006. 2, 27

[7] S. Sabanovic, Marek Piotr Michalowski, and Reid Simmons. **Robots in the wild: observing human-robot social interaction outside the lab**. In *Proceedings of the 9th International Workshop on Advanced Motion Control (AMC 2006)*, pages 596 – 601. IEEE, March 2006. 2, 14, 16

[8] Cynthia Breazeal. **Towards Sociable Robots**. *Robotics and Autonomous Systems*, **42**, 2002. 7

[9] Laurel D. Riek. **Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines**. *Journal of Human-Robot Interacion*, 2012. 8

[10] Aaron Steinfeld, Odest Chadwicke Jenkins, and Brian Scassellati. **The Oz of Wizard: Simulating the Human for Interaction Research**. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, March 2009. 8

[11] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita. **Interactive Humanoid Robots for a Science Museum**. *Intelligent Systems, IEEE*, **22**(2):25–32, 2007. 8, 12, 46

[12] Helge Huettenrauch, Kerstin S. Eklundh, Anders Green, and Elin. **Investigating Spatial Relationships in Human-Robot Interaction**. In *International Conference on Intelligent Robots and Systems*, pages 5052–5059, 2006. 8, 15, 17, 19, 20

[13] Takayuki Kanda, Takayuki Hirano, Daniel Eaton, and Hiroshi Ishiguro. **Interactive Robots As Social Partners and Peer Tutors for Children: A Field Trial**. *Hum.-Comput. Interact.*, **19**(1):61–84, June 2004. 8

[14] Michael Montemerlo, Joelle Pineau, Nicholas Roy, Sebastian Thrun, and Vandi Verma. **Experiences with a Mobile Robotic Guide for the Elderly**. In *Eighteenth National Conference on Artificial Intelligence*, pages 587–592, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence. 8

[15] Roland Siegwart, Kai O. Arras, Samir Bouabdallah, Daniel Burnier, Gilles Froidevaux, Xavier Greppin, Björn Jensen, Antoine Lorotte, Laetitia Mayor, Mathieu Meisser, Roland Philippsen, Ralph Piguet, Guy Ramel, Gregoire Terrien, and Nicola Tomatis. **Robox at Expo.02: A large-scale installation of personal robots**. *Robotics and Autonomous Systems*, **42**(3–4):203 – 222, 2003. Socially Interactive Robots. 8

[16] Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk Hahnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. **The Interactive Museum Tour-Guide Robot**. In *Proc. AAAI/IAAI*, pages 11–18, 1998. 8, 17, 46

[17] Kunz-C. & Nourbakhsh I. R. Willeke, T. **The History of the Mobot Museum Robot Series: An Evolutionary Study**. In *FLAIRS Conference*, pages 514–518, 2001. 8, 11, 47

[18] Kerstin Dautenhahn. **Socially intelligent robots: dimensions of human-robot interaction**. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362**(1480):679–704, 2007. 9

[19] Caryl E. Rusbult and Paul A. M. Van Lange. **Interdependence, Interaction, and Relationships**. *Annual Review of Psychology*, **54**(1):351–375, 2003. 9

[20] Bob G. Witmer and Michael J. Singer. **Measuring Presence in Virtual Environments: A Presence Questionnaire**. *Presence: Teleoper. Virtual Environ.*, **7**(3):225–240, June 1998. 9

[21] Matthew Lombard and Theresa Ditton. **At the Heart of It All: The Concept of Presence.** *J. Computer-Mediated Communication*, **3**(2):0, 1997. 9

[22] A. Keay and M. Graduand. **Emergent phenomena of robot competitions: Robot identity construction and naming**. In *Advanced Robotics and its Social Impacts (ARSO), 2011 IEEE Workshop on*, pages 12–15, Oct 2011. 10

[23] Mark Coeckelbergh. **You, robot: on the linguistic construction of artificial others**. *AI & SOCIETY*, **26**(1):61–69, 2010. 10

[24] Wolfram Burgard, Armin B. Cremers, Dieter Fox, Dirk Hahnel, Gerhard Lakemeyer, Dirk Schulz, Walter Steiner, and Sebastian Thrun. **Experiences with an Interactive Museum Tour-Guide Robot**. *Artificial Intelligence*, **114**(1):3 – 55, 1999. 10

[25] Sebastian Thrun, Maren Bennewitz, Wolfram Burgard, Armin B. Cremers, Frank Dellaert, Dieter Fox, Dirk Hähnel, Charles R. Rosenberg, Nicholas Roy, Jamieson Schulte, and Dirk Schulz. **MINERVA: A Second-Generation Museum Tour-Guide Robot**. In *1999 IEEE International Conference on Robotics and Automation, Marriott Hotel, Renaissance Center, Detroit, Michigan, May 10-15, 1999, Proceedings*, pages 1999–2005, 1999. 10

[26] Daphne Karreman, Geke Ludden, and Vanessa Evers. **Visiting Cultural Heritage with a Tour Guide Robot: A User Evaluation Study in-the-Wild**. In Adriana Tapus, Elisabeth André, Jean-Claude Martin, François Ferland, and Mehdi Ammi, editors, *Social Robotics*, pages 317–326, Cham, 2015. Springer International Publishing. 12

[27] Catherina R. Burghart and Aaron Steinfeld. **Proceedings of Metrics for Human-Robot Interaction, a Workshop at ACM/IEEE HRI 2008**, March 2008. 13

[28] Luca Brayda and Ryad Chellali. **Measuring Human-Robots Interactions**. *International Journal of Social Robotics*, pages 1–3, 2012. 13

[29] 2003 Performance Metrics for Intelligent Systems (PerMIS 03) Workshop. *Metrics for Evaluating Human-Robot Interactions*, 2003. 13

[30] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. **Common Metrics for Human-Robot Interaction**. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, HRI '06, pages 33–40, New York, NY, USA, 2006. ACM. 13, 16

[31] Astrid Weiss, Regina Bernhaupt, Michael Lankes, and Manfred Tscheligi. **The usus evaluation framework for human-robot interaction**. In *in AISB2009: Proceedings of the Symposium on New Frontiers in Human-Robot Interaction*, 2009. 14, 15, 16

[32] Christoph Bartneck, Dana Kuli, Elizabeth Croft, and Susana Zoghbi. **Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots**. *International Journal of Social Robotics*, **1**(1):71–81, 2009. 14, 16, 142

[33] James E. Young, JaYoung Sung, Amy Voida, Ehud Sharlin, Takeo Igarashi, Henrik I. Christensen, and Rebecca E. Grinter. **Evaluating Human-Robot Interaction**. *International Journal of Social Robotics*, **3**(1):53–67, 2010. 14, 16

[34] Ross Mead, Amin Atrash, and Maja J. Matarić. **Proxemic feature recognition for interactive robots: automating metrics from the social sciences**. In *Proceedings of the Third international conference on Social Robotics*, ICSR'11, pages 52–61, Berlin, Heidelberg, 2011. Springer-Verlag. 15, 16, 116

[35] Paul Marshall, Yvonne Rogers, and Nadia Pantidi. **Using F-formations to analyse spatial patterns of interaction in physical environments**. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, CSCW '11, pages 445–454, New York, NY, USA, 2011. ACM. 17, 18, 20

[36] Edward Hall. *The hidden dimension*. The Bodley Head Ltd, London, UK, 1966. 17

[37] Adam Kendon. **Spacing and Orientation in Copresent Interaction**. In Anna Esposito, Nick Campbell, Carl Vogel, Amir Hussain, and Anton Nijholt, editors, *Development of Multimodal Interfaces: Active Listening and Synchrony*, **5967** of *Lecture Notes in Computer Science*, pages 1–15. Springer Berlin Heidelberg, 2010. 18, 176, 177, 178, 179, 180, 181, 189

[38] Elena Torta, Raymond H. Cuijpers, James F. Juola, and David Van Der Pol. **Modeling and testing proxemic behavior for humanoid robots**. *International Journal of Humanoid Robotics*, **9**(4), DEC 2012. 19

[39] H. Kuzuoka, Y. Suzuki, J. Yamashita, and K. Yamazaki. **Reconfiguring spatial formation arrangement by robot body orientation**. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on*, pages 285–292, 2010. 19

[40] Anais Garrell and Alberto Sanfeliu. **Cooperative social robots to accompany groups of people**. *Int. J. Rob. Res.*, **31**(13):1675–1701, Nov 2012. 19

[41] Marco Costa. **Interpersonal Distances in Group Walking**. *Journal of Nonverbal Behavior*, **34**(1):15–26, 2010. 19, 20, 51

[42] S. Bandini, A. Gorrini, L. Manenti, and G. Vizzari. **Crowd and Pedestrian Dynamics: Empirical Investigation and Simulation**. In *Proceedings of Measuring Behavior*, 2012. 20, 21, 40

[43] Andrea Bauer, Dirk Wollherr, and Martin Buss. **HUMAN-ROBOT COLLABORATION: A SURVEY**. *International Journal of Humanoid Robotics*, **05**(01):47–66, 2008. 20

[44] Photchara Ratsamee, Yasushi Mae, Kenichi Ohara, Tomohito Takubo, and Tatsuo Arai. **HUMAN–ROBOT COLLISION AVOIDANCE USING A MODIFIED SOCIAL FORCE MODEL WITH BODY POSE AND FACE ORIENTATION**. *International Journal of Humanoid Robotics*, **10**(01):1350008, 2013. 20

[45] P. Trautman and A Krause. **Unfreezing the robot: Navigation in dense, interacting crowds**. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 797–803, Oct 2010. 20, 46

[46] Stefania Bandini, Andrea Gorrini, and Giuseppe Vizzari. **Towards an integrated approach to crowd analysis and crowd synthesis: A case study and first results**. *Pattern Recognition Letters*, **44**(0):16 – 29, 2014. Pattern Recognition and Crowd Analysis. 21, 51

[47] Aryel Beck, Lola Canamero, Antoine Hiolle, Luisa Damiano, Piero Cosi, Fabio Tesser, and Giacomo Sommavilla. **Interpretation of Emotional Body Language Displayed by a Humanoid Robot: A Case Study with Children**. *International Journal of Social Robotics*, **5**(3):325–334, 2013. 27

[48] CosmoCaixa. **Datos generales de CosmoCaixa Barcelona**, 2013. [Online; accessed 16-May-2013]. 29

[49] P. Kline. *The Handbook of Psychological Testing*. Routledge, 2000. 65, 143, 153

[50] Carl F. DiSalvo, Francine Gemperle, Jodi Forlizzi, and Sara Kiesler. **All Robots Are Not Created Equal: The Design and Perception of Humanoid Robot Heads**. In *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '02, pages 321–326, New York, NY, USA, 2002. ACM. 99

[51] Yeow Kee Mirnig, Nicoleand Tan, Tai Wen Chang, Yuanwei Chua, Tran Anh Dung, Haizhou Li, and Manfred Tscheligi. **Screen Feedback in Human-Robot Interaction: How to Enhance Robot Expressiveness**. *The 23rd IEEE International Symposium on Robot and Human Interactive Communication IEEE RO-MAN 2014*, August 2014. 99

[52] Owen Churches, Mike Nicholls, Myra Thiessen, Mark Kohler, and Hannah Keage. **Emoticons in mind: An event-related potential study**. *Social Neuroscience*, **9**(2):196–202, 2014. PMID: 24387045. 99

[53] Mohammed Moshiul Hoque, Quazi Delwar Hossian, and Kaushik Deb. **Effects of Robotic Blinking Behavior for Making Eye Contact with Humans**. In Malay Kumar Kundu, Durga Prasad Mohapatra, Amit Konar, and Aruna Chakraborty, editors, *Advanced Computing, Networking and Informatics- Volume 1*, pages 621–628, Cham, 2014. Springer International Publishing. 107

[54] Marta Díaz-Boladeras, Dennys Paillacho, Cecilio Angulo, Oriol Torres, Jonathan Gonzalez, and Jordi Albo-Canals. **Evaluating Group-Robot Interaction in Crowded Public Spaces: A Week-Long Exploratory Study in the Wild with a Humanoid Robot Guiding Visitors Through a Science Museum**. *International Journal of Humanoid Robotics*, **12**(4):1550022, Mar 2015. 134, 189

[55] C. Dondrup, N. Bellotto, and M. Hanheide. **Social distance augmented qualitative trajectory calculus for Human-Robot Spatial Interaction**. In *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*, pages 519–524, Aug 2014. 175, 177

[56] M. Hanheide, A. Peters, and N. Bellotto. **Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus**. In *RO-MAN, 2012 IEEE*, pages 689–694, Sept 2012. 175

[57] Iolanda Leite, Marissa McCoy, Daniel Ullman, Nicole Salomons, and Brian Scassellati. **Comparing Models of Disengagement in Individual and Group Interactions**. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15, pages 99–105, New York, NY, USA, 2015. ACM. 176

[58] Anthony G. Cohn and Jochen Renz. **Chapter 13 Qualitative Spatial Representation and Reasoning**. In Vladimir Lifschitz Frank van Harmelen and Bruce Porter, editors, *Handbook of Knowledge Representation*, **3** of *Foundations of Artificial Intelligence*, pages 551 – 596. Elsevier, 2008. 177

[59] C. Dondrup, N. Bellotto, and M. Hanheide. **A Probabilistic Model of Human-Robot Spatial Interaction Using a Qualitative Trajectory Calculus**. In *AAAI Spring Symposium – Qualitative Representations for Robots*, pages 18–25, 2014. 177

[60] Nicola Bellotto, Marc Hanheide, and Nico Van de Weghe. **Qualitative Design and Implementation of Human-Robot Spatial Interactions**. In Guido Herrmann, MartinJ. Pearson, Alexander Lenz, Paul Bremner, Adam Spiers, and Ute Leonards, editors, *Social Robotics*, **8239** of *Lecture Notes in Computer Science*, pages 331–340. Springer International Publishing, 2013. 177

[61] Marc Hanheide, Annika Peters, and Nicola Bellotto. **Analysis of human-robot spatial behaviour applying a qualitative trajectory calculus**. In *RO-MAN, 2012 IEEE*, pages 689–694, Sept 2012. 177

[62] Frank Dylla, Martin Coors, and Mehul Bhatt. **Socially Compliant Navigation in Crowded Environments**. In *Spatial and Temporal Dynamics (STeDy) Workshop at ECAI 2012*, pages 9 – 17, 2012. 177, 189

[63] Frank Dylla, Arne Kreutzmann, and Diedrich Wolter. **A Qualitative Representation of Social Conventions for Application in Robotics**. In *2014 AAAI Spring Symposium - Qualitative Representations for Robots*, number SS-14-06 in Papers from the 2014 AAAI Spring Symposium, 2014. 177

[64] Christian Goerick. **Towards Cognitive Robotics**. In Bernhard Sendhoff, Edgar Körner, Olaf Sporns, Helge Ritter, and Kenji Doya, editors, *Creating Brain-Like Intelligence: From Basic Principles to Complex Intelligent Systems*, chapter Towards Cognitive Robotics, pages 192–214. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. 178

[65] Alessandro Couyoumdjian, Francesco Di Nocera, and Fabio Ferlazzo. **Functional representation of 3D space in endogenous attention shifts**. *The Quarterly Journal of Experimental Psychology Section A*, **56**(1):155–183, 2003. PMID: 12587900. 178

[66] Rory J. Bufacchi, Meng Liang, Lewis D. Griffin, and Giandomenico D. Iannetti. **A geometric model of defensive peripersonal space**. *Journal of Neurophysiology*, **115**:218 – 225, 2016. 179

[67] Elena Pacchierotti, Henrik I. Christensen, and P. Jensfelt. **Embodied social interaction for service robots in hallway environments**. In *Proc. of the International Conference on Field and Service Robotics (FSR'05)*, July 2005. 189

[68] Pablo Bustos, Luis Jesús Manso, Juan Pedro Bandera Rubio, Adrián Romero-Garcés, Luis V. Calderita, Rebeca Marfil, and Antonio Bandera. **A Unified Internal Representation of the Outer World for Social Robotics**. In *Robot 2015: Second Iberian Robotics Conference - Advances in Robotics, Lisbon, Portugal, 19-21 November 2015, Volume 2*, pages 733–744, 2015. 189