Lessons from the MOnarCH Project

João Silva Sequeira[1] and Isabel Aldinhas Ferreira[2]
[1]*IST/ISR, Universidade de Lisboa, Lisbon, Portugal*
[2]*Centro de Filosofia da Universidade de Lisboa, Universidade de Lisboa, Lisbon, Portugal*

Keywords: Human-Robot Interaction, Social Robotics.

Abstract: The paper describes the main conclusions issued from three years of development, of which approximately eighteen months of observations, of an autonomous social robot interacting with children and adults in the Pediatrics ward of an Oncological hospital. After this long period of trials and continuous interaction, the integration of the robot in that particular social environment can be considered highly successful. The results taken from all the long run experiments yield valuable lessons in what relates social acceptance and user experience to be considered in the case of robot deployment at institutions or even at households. These are detailed in the paper.

1 INTRODUCTION

The main goal of the MOnarCH project[1] is to test the integration of robots in a social environment, assessing the relationships established between humans and robots and identifying guidelines for developments to be incorporated in future social robots.

The project is evolving in the Pediatrics ward on an Oncological hospital[2] where the robot engages in edutainment activities with the inpatient children. Though this environment poses no special difficulties on the locomotion of a wheeled robot, it can be challenging in what concerns obstacle avoidance as the inhabitant children can use the space of the ward to play with bulky toys and tend to be naturally entropic.

The robot is shown in Figure 1. It is about the height of an 8-10 year old child, capable of omnidirectional motion with linear velocities up to 2.5 m/s, thus capable to walk along a person at fast pace. The volume of the robot is compatible to that of a child companion.

Figure 1: The MOnarCH robot (source: SIC TV Network).

There are a number of factors difficult to quantize when dealing with inpatient children. Nowadays there is a wide variety of toy robots commercially available, some with sophisticated interfacing skills. Nevertheless, the great majority of children had never been close to a robot this size. Even though many of them already know what a robot is and have clear prototypical images of a robot (see (Sequeira and Ferreira, 2014)) this is a perception obtained from movies, books, toys, etc. At the hospital it has been observed that children often show surprise when they see the robot for the first time moving autonomously. This initial perception may have a lasting effect and influence future behaviors (see for instance (Boden, 2014)).

This paper does not report results supported on rigid observation methodologies as often used in the literature, with Likert questionnaires and statistical processing of the results (preliminary analysis can be found in project MOnarCH Deliverables, (MOnarCH Project Consortium, 2015; MOnarCH Project Con-

---

[1] www.monarch-fp7.eu

[2] IPOLFG, in Lisbon.

Sequeira, J. and Ferreira, I.
Lessons from the MOnarCH Project.
DOI: 10.5220/0005998102410248
In *Proceedings of the 13th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2016) - Volume 1*, pages 241-248
ISBN: 978-989-758-198-4

sortium, 2016)). The main reason is that such approach can disturb the normal operation of the environment (the Pediatrics ward in this case) and bias any results (the argument in (Boden, 2014) on the importance of initial perceptions can also be used here). Instead, the paper is supported on direct observations of all the actors in the environment. The paper's conclusions are grounded on almost 18 months of intense observation of the reactions of children and adults at the hospital with the robot evolving through different stages of development. In a sense, it can be assumed that multiple inpatient children with essentially similar social behaviors were interacting, directly and indirectly, with the robot. This number is estimated in the order of hundreds. Therefore the number of children that knew about the robot is large enough for the observations to have some significance.

We claim that social experiments involving robots must unfold with the dynamics of the social environment itself. Otherwise there is a risk of biasing the results. Rigid planning of experiments, as often done in lab environments often cannot be used. In fact, while preparing an experiment a team is changing the environment and potentially biasing the reaction of the inhabitants.

The paper is organized as a collection of the main lessons learned, one per section, without any ranking concerns. All the lessons are equally important. Even though some of these lessons are well known of Robotics developers, it is worth to recall them. Moreover, these lessons are not intended to be a methodological approach to project development.

2 DESIGNING THE EXPERIMENTS

The hospital environment is socially constrained often to a high degree. Besides the ethical and legal regulations, there are a large number of unwritten rules (the social rules) the robot should comply with to maximize the acceptance by the inhabitants of the social environment, i.e., children, parents, relatives and visitors, and staff. Moreover, some of these rules may change in time depending, for example, on the occupancy of the ward.

The routines of the hospital cannot be disturbed at any time and even though the robot is constantly under development, experiments must occur only when the adequate conditions in the social environment emerge. This means that seldom the experiments can be planned. Instead, the robot must match its capabilities to the current state of the environment and generate observations from which relevant conclusions can be withdrawn. In a sense, experiments must occur as the (uncontrolled) events arise. This randomness allowed nonetheless a rough scheduling for classes of experiments.

The robot design followed principles also recommended by other authors, (see for example (Kim, 2007)) namely that of using test frameworks and ensure that the project team have a thorough understanding of the HRI interfaces, of the environment, and of the influence of design aspects. In MOnarCH, besides the testing in the hospital, numerous tests of isolated components were conducted i lab testbeds at some of the partners of the project.

The robot was first deployed at the environment mainly in a static exhibition, most of the time being kept inside a room but being available for the children to see upon request to the staff. This first period lasted for around 5 months, during which the robot was taken out in order for the team to assess different functionalities, e.g., the navigation system.

The following period lasted for approximately 7 months during which the core functionalities of the system were developed, namely the games the robot can play with the children and the communicative actions and expressions used for human-robot interaction. The robot was shown in a normal development cycle (sometimes failing) and the children could watch it during the test runs.

During the last period the robot was used in semi and full autonomy, wandering around the ward, playing with the children in some occasions. The tuning of the communicative actions and expressions is currently taking the majority of the effort. This is a lengthy process, requiring the careful observation of the environment and trial runs to assess the efficacy of the interactions.

The Wizard-of-Oz (WoZ) technique (Green et al., 2008) has proved highly valuable in early stages, namely as it allowed the testing of a number of interactions (communicative acts in the project parlance, see (Alonso-Martin et al., 2013)) for later use. In these experiments, a Wizard controlled some of the verbal capabilities of the robot and could decide navigation goals. The Wizard role was played by a project team member located remotely but maintaining a good perception of the environment. This form of WoZ usage, with real technology in a real environment, has been reported to be seldom used in research, (Riek, 2012).

3 SOFTWARE

The underlying structure of MOnarCH is a network of software components. It can be assumed that each of them was developed independently and that the communication among them is based on some predefined protocol. The dynamics of each of the software components is of major importance to the overall stability and performance. A component that does not clearly broadcast its state can prevent other components to run properly. Also, the handling of exception conditions must be carefully defined, e.g., transitions between states must occur cleanly. In a sense this amount to “controllability”and “observability” concerns on the overall system (the relevance of which has been recognized by some authors, e.g., (Naganathan and Eugene, 2009)).

The visible part of the system is composed by the outer shell of the robot (Figure 1), the corresponding Human-Robot Interaction (HRI) interfaces, and sensors. The HRI interfaces must be used much in a human way, meaning that visual effects must be well coordinated with sound effects so that the observers get the correct perception and the interaction is efficient. The MOnarCH robot has two 1-dof arms and a 1-dof neck, variable luminosity eyes and cheeks, a LED matrix for a mouth, loudspeakers, microphone, bumpers, touch sensors placed in strategic places of the body shell, and a RFID tag reader. The adequate coordination of HRI interfaces can be framed as a stability problem.

In MOnarCH, a key architectural feature that accounts for this concern is called Situational Awareness Module (SAM). In a sense this is a routing information matrix that defines what are the information exchanges, and, if necessary, applies basic transformations to that information (see (Messias et al., 2014)).

Moreover, performance and stability concerns have also been addressed through the inclusion of watchdog mechanisms (see (MOnarCH Project Consortium, 2016)).

The ROS middleware (www.ros.org) supports the integration of all components. Key features of the development are (i) local networking capabilities with distributed node representation, (ii) global variables accessible from any node in the system (the topics in the ROS middleware), (iii) global networking is useful for monitoring, (iv) fully independent control of each component, (v) callback mechanisms to deal with asynchronous events (actionlib), and (vi) state machine frameworks (smach) for fast behavior development, among others. Structures like the SAM module can be easily implemented in the ROS middleware using the built-in facilities. In a sense, a SAM is a collection of direct mappings between ROS topics.

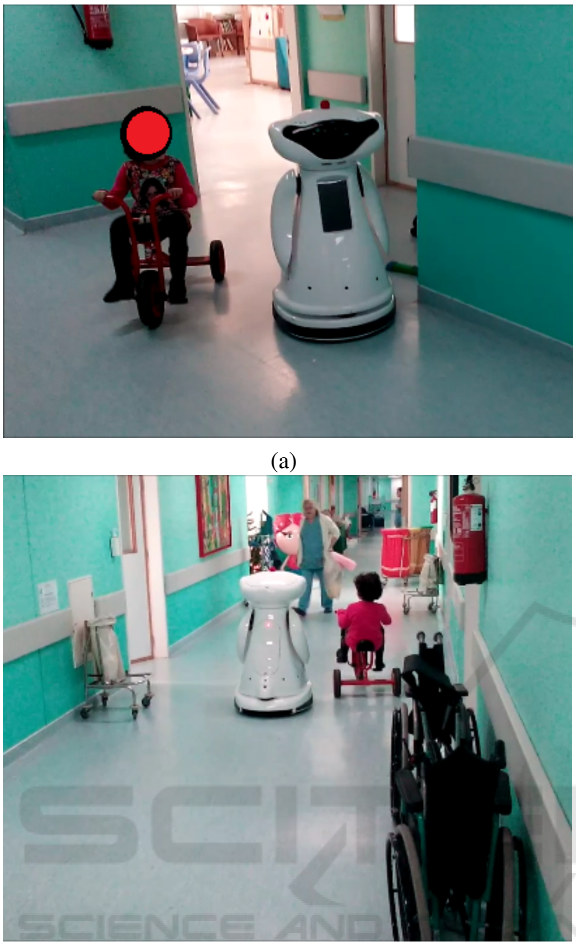
4 MOTION

Motion is the most basic tool for a robot to interact with humans. In the particular case of inpatient children (in the Pediatrics age range velocity and acceleration must be carefully selected. Relatively small accelerations can be perceived by children as menacing. Similarly, fast decelerations can be perceived as if the robot is untrusty.

Obstacle avoidance is tightly connected to motion. Real environments can be really challenging, e.g., children can move in unpredictable ways (see Figure 2), and collisions are to be expected. Whenever a robot collides it is a good practice that it clearly shows that it knows a collision just occurred. A practical way is to issue a non-verbal sound that can easily be associated to a collision, e.g., ouch.

The space in the Pediatrics ward is reasonably structured (see the pictures presented in the paper). Using an a priori map built with the gmapping ROS package, (http://wiki.ros.org/gmapping, 2016), is obtained by teleoperating the robot such that all the ward is covered it is possible to localize the robot only with the information from the laser scans. The navigation system in MOnarCH is based on a potential field approach combining a fast marching method to construct an optimal path to the goal and the repulsive fields from obstacles, (Ventura and Ahmad, 2014). Localization is based on the AMCL package, available in ROS, (http://wiki.ros.org/amcl, 2016), and uses information from two laser range finders that scan the full space around the robot. The accuracy of the pose estimate is in the order of a few cm and degrees and the system has been show to have a remarkable robustness. Observations so far indicate that, on average, the robot loses its localization less than once a day. When this happens the unusual behavior of the robot is enough for an external observer to recognize it and either alert the project team or push the robot to a specific location where it can auto-localize easily.

In addition to the motion of the body of the robot, the 1-dof neck can rotate right/left and it conveys quite effectively the idea that the robot has a clear focus of attention. If the head turns because a person is detected on that side a child tends to conclude that the robot knows where the person is. If the head turns to the opposite side where a person is then a plausible perception is that the robot does not care about the person and has some other focus of attention. In both cases the observations suggest that the motion of the



(a)

(b)

Figure 2: Children running a toy bike chasing the robot; the robot must account as much as possible for sudden deviations of the straight line path by the child.

neck is a relevant feature for human-robot interaction.

Also, the two 1-dof arms are used to improve liveness. A simple balancing of the arms while the robot is moving can easily convey the perception that the robot has a focus while moving or a goal location where to go.

5 HUMAN-ROBOT INTERACTION INTERFACES

A basic speech interface is used to convey either verbal or non-verbal sounds. Tone and volume are highly relevant for a sound message to be perceived properly.

Short verbal and non-verbal sounds can be used to show liveness. The trigger of such sounds can be done following an exponentially distributed process, with the corresponding parameter eventually tuned to match the environment.

Experiments performed with off-the-shelf Text-To-Speech systems using adult voices have been shown to trigger negative reactions by some children, namely of very young age (even if the sentences are not aggressive). As an alternative, pre-recorded childish tone speech has been very well received both by children and adults.

Off-the-shelf automatic speech recognition (ASR) has been also tested (using for example the SpeechRecognition-3.4.6 Python based package) Reasonable recognition rates (from the perspective of basic HRI) can be achieved when sentences are spoken with clear diction which is often not the case with children. This reduces the range of interactions that an be supported on ASR.

A touchscreen in the chest is useful to (i) quickly convey information on the “emotional state” of the robot through emoticon-like graphics. Children tend to be curious about the meaning of the graphics shown and quickly come up with their own interpretation, which in a sense amounts to the establishing relation with the robot.

Touch sensors are placed in various locations below the outer shell. These are non-visible interfaces. This type of interface tends to be effective with young children as they are more prone to engaging in touching behaviors.

Facial expressions, generated by the LED mouth, eyes, and cheeks, tend to be interpreted as indicators that the robot is active (often in conjunction with the graphics displayed in the chest screen).

6 CHILDREN’S REACTIONS

Though the Pediatrics ward covers the full range of Pediatric ages, the observation showed that ages in the range 3-8 are the most interested. Moreover, children with mild degrees of Asperger syndrome are also highly receptive to robot interactions. Also, children love to be recognized by name by the robot. Wearing a flexible RFID tag attached to the clothes allows the robot to recognize the children and greet them occasionally.

Children above 10-12 years old may look at the robot but seldom initiate an interaction. They often direct gaze to the robot but rarely touch it.

During the initial stages of the project the robot was deployed at the hospital and when no team member was around it was kept in a separate room, with reserved access. The inpatient children know the reserved location of robot and often ask the staff to go there to see it, just to see the robot sleeping.

Currently the robot is kept in a public space, that is, a docking station that is also used to charge batteries. During the periods at the docking station (the sleeping periods) the robot maintains a small degree of liveness, reacting to touch stimulae. Even in this almost static display children often look for the robot either to just see it or to touch it and observe the reaction induced by the liveness module.

7 TEAM MEMBERS

The team members working at the environment must be integrate themselves socially to avoid becoming biasing/disturbing/distracting factors. Successful integration smooths potential social barriers that some people may rise against the robot. In fact this strategy resulted from a hard constraint imposed by the hospital from the very beginning of the project, i.e., the hospital should continue to operate normally and not be transformed in a robotics lab, with the team members disturbing the ward.

The integration of the team members occurred through multiple forms, e.g., explaining the robot to parents/visitors/staff/children and bringing the robot to the rooms upon requests by inpatient children with restricted mobility (see figure 3).

8 PARENTS AND VISITORS REACTIONS

Parents and visitors were informed about the project through postcards visible in the walls of the Pediatric ward. The observations clearly show that parents are curious about the robot. Direct interactions do occur, with parents and visitors often looking at the robot while moving in the ward, take pictures, and even touching the robot.

As the autonomy of the robot increased parents and visitors started to become confident asking the team to adjust the behaviors of the robot. A typical example is the adjustment of the snoring behavior the robot exhibits when sleeping, or in the non-verbal sounds the robot occasionally issues.

9 STAFF'S REACTIONS

Staff was given initially basic information on the project and on the robot, and on actions to adopt in case of the robot is disturbing the normal operation of the ward. A basic instruction was that the robot has

(a)

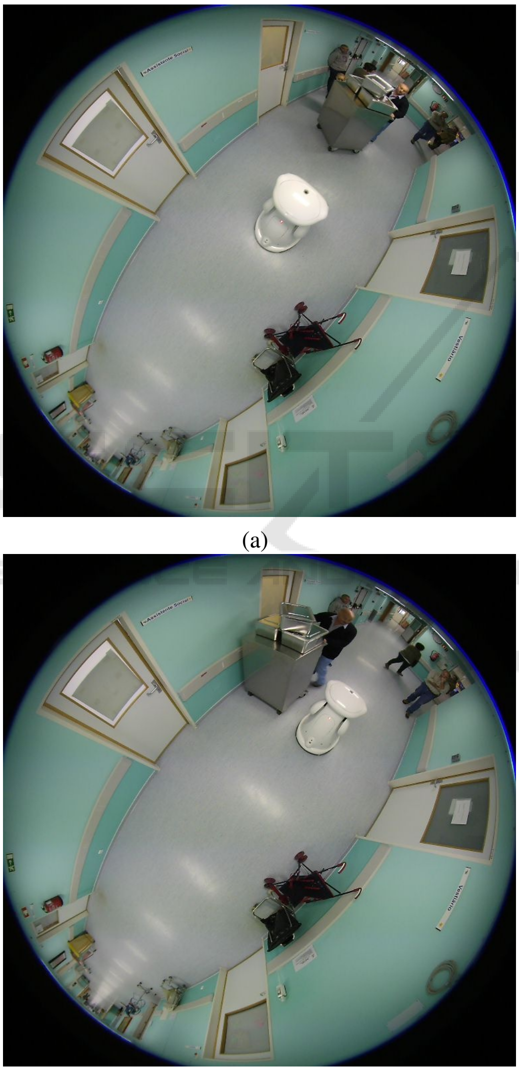
(b)

(c)

Figure 3: Team members and hospital staff assisting children during interactions.

an emergency button that can be pushed at any time to stop it and allow any staff member to push it away. During the whole period no observations of any staff member pushing the button were made.

In fact, the staff adapted to this additional member that is often seen moving along the main corridor of the Pediatrics ward. Figure 4 shows a staff member pushing a cart and deviation from the robot. The staff knows that the robot is capable of deviating from static and moving obstacles but does not wait and performs the evasive maneuver


(a)

(b)

Figure 4: Staff adapting to the behavior of the robot (acting compliantly).

In some occasions it has been observed that some staff members like to test the obstacle avoidance capabilities of the robot, usually for the fun of it, and try to verify the conditions under which the robot gets trapped.

Similarly to children, a number of staff members appreciates being recognized by the robot. RFID tags worn in the service uniforms allow the robot to occasionally greet the members.

10 METRICS AND ASSESSMENT

A number of metrics is being used to assess the project. These are based on the activation rates of selected micro-behaviors (units of activations per second), annotated from the observations of the experiments (this is a common strategy used by multiple authors, e.g., (Mead et al., 2011; Scassellati, 2005; Dautenhahn and Werry, 2002).

Table 1 shows results of a recent experiment at the hospital. The experiment took place over a period of two consecutive days.

Table 1: Activation rates ($\times 10^{-2}$) for a typical experiment (total duration of 5685 seconds over 2 consecutive days and a total of 249 micro-behaviors annotated, units are activations per second).

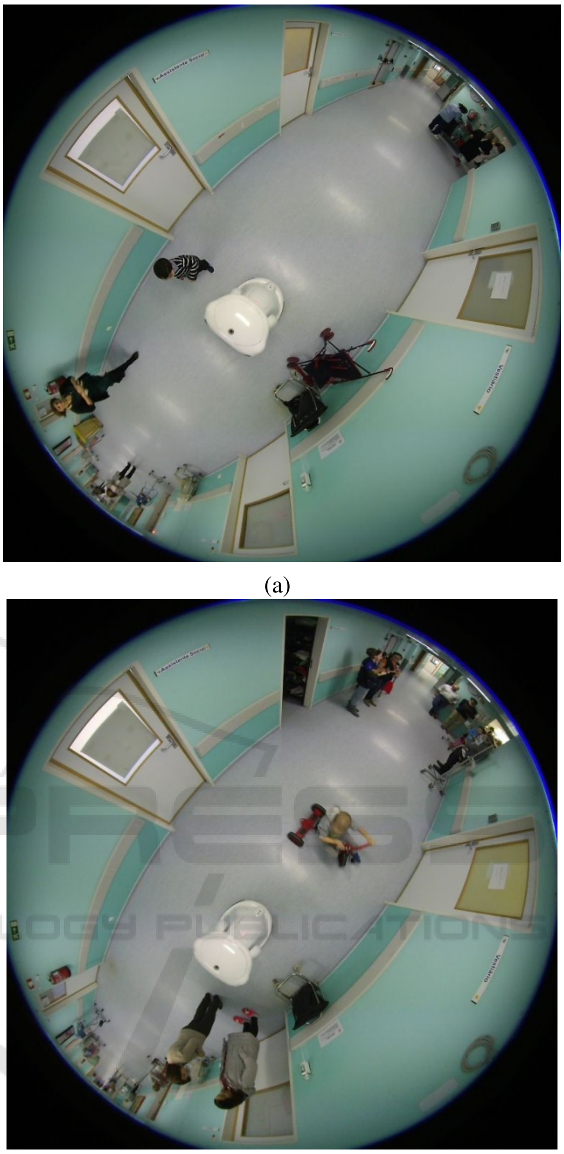
Micro-behaviors	Activation rates (robot showing low liveness)	Activation rates (robot showing high liveness)
1. Looking towards the robot, without moving	1.0	0.86
2. Looking towards the robot and moving (around, ahead and/or at the back of the robot)	0.64	0.78
3. Touching the robot	0.16	0.16
4. Aggressive movement towards the robot	0.03	0
5. Ignoring the robot	1.9	1.1
6. Following the robot	0.32	0.23
7. Compliant behavior towards the robot	0.61	0.82

The values emphasize the importance of using high levels of liveness. For example, the values suggest that an intention to initiate an interaction may be bigger if the robot shows that it is alive (lines 2 and 5). Also, if the robot shows aliveness aggressive micro-behaviors may be less likely (line 3). The values in line 6 can be interpreted as a result of curiosity to understand the behavior of the robot when the liveness shown is not enough to decide about interacting with it.

Complex metrics can be defined over the activation rates of collections of micro-behaviors. For example, the acceptance of the robot can be measured by comparing a combination of micro-behaviors 2, 3, 6, 7 against a combination of micro-behaviors 1, 4, 5. Designing these combinations will be in general somewhat subjective but it may provide nonetheless useful guidance to adjust the behaviors of the robot (see (MOnarCH Project Consortium, 2015; MOnarCH Project Consortium, 2016) for additional details).

Manual annotations are often subject to large errors. To achieve statistical significance the annotations must (i) focus on experiments made in similar conditions, namely in what concerns the robot and the characteristics of the population interacting with it, and (ii) annotated by multiple persons (or multiple systems in case of automatic annotation). This is often difficult due to the constantly changing of the inpatient population. For example, there are periods where most of the children is in isolation rooms (and hence cannot interact with the robot) and periods in which many children are allowed to leave the rooms (thus being able to play with the robot). This means that either the experiments take place in close days, or it may be necessary to wait for the right conditions to be present.

Independently of the details on how the activation rates are combined, it seems clear that activation rates provide interesting quantitative information. The use of automated system to annotate the micro-behaviors could further provide, for example, the time between activations of micro-behaviors and their duration, which are necessary for a complete study of their probabilistic nature. The use of such systems, e.g., (www.noldus.com, 2016), is out of the possibilities of the project.



(a)

(b)

Figure 5: Snapshots acquired during the liveness experiments.

11 CONCLUSIONS

It is commonly accepted that a real social environment is substantially different from common laboratory environments. Achieving successful integration of a robot in a real social environment, such as the Pediatrics ward of the MOnarCH project has been shown to depend on multiple factors, some outside the pure engineering scope. Moreover, current off-the-shelf and state-of-the-art robotics seem well adapted to the objectives of the project, though technological improvements may simplify similar systems. There is no evidence that the lessons learned depend on the specific type of robot used (wheeled, omnidirectional) and hence may be useful in projects involving other types of social robots.

Overall, the high acceptance of the robot by everyone at the hospital is a clear indicator that the strategy that led to the lessons in the paper was successful.

A quantitative assessment in medical terms of a robot such as the one developed in the MOnarCH project may take several years and hence it out of the scope of the project. Meanwhile, all the indicators observed in MOnarCH suggest that social robots may have a highly positive role in socially difficult environments.

ACKNOWLEDGEMENTS

Work supported by projects FP7-ICT-9-2011-601033 (MOnarCH) and FCT [UID/EEA/50009/2013].

REFERENCES

Alonso-Martin, F., Malfaz, M., Sequeira, J., Gorostiza, J., and Salichs, M. (2013). A Multimodel Emotion Detection System during Human-Robot Interaction. *Sensors Journal*, 13:15549–15581.

Boden, M. (2014). Aaron Sloman: A Bright Tile in AI's Mosaic. In Jeremy L. Wyatt and Dean D. Petters and David C. Hogg, editor, *From Animals to Robots and Back: Reflections on Hard Problems in the Study of Cognition*. Springer. Cognitive Systems Monographs 22.

Dautenhahn, K. and Werry, I. (2002). A Quantitative Technique for Analysing Robot-Human Interactions. In *Procs. of the 2002 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*. Switzerland.

Green, S., Richardson, S., Stiles, R., Billinghurst, M., and Chase, J. (2008). Multimodal Metric Study for Human-Robot Collaboration. In *Procs. 1st Int. Conf. on Advances in Computer-Human Interaction*, pages 218–233. Sainte-Luce, Martinique, 10-15 February.

http://wiki.ros.org/amcl (2016). Adaptive Monte-Carlo Localization. Accessed April 2016.

http://wiki.ros.org/gmapping (2016). Gmapping. Accessed April 2016.

Kim, M. (2007). The HRI Experiment Framework for Designer. In *Procs. of The 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2007)*. Jeju, Korea, 26-29 August.

Mead, R., Atrash, A., and Mataric, M. (2011). Proxemic Feature Recognition for Interactive Robots: Automating Metrics from Social Sciences. In *In Social Robotics*, volume 7072 of *Lecture Notes in Computer Science*, pages 52–61. Springer.

Messias, J., Ventura, R., and Lima, P. (2014). The MOnarCH Situation Awareness Module. Project MOnarCH deliverable document available at http://www.monarch-fp7.eu.

MOnarCH Project Consortium (2015). Project Deliverable D8.8.4 - Long-Run MOnarCH Experiments at IPOL. http://users.isr.ist.utl.pt/∼jseq/MOnarCH/Deliverables/D8.8.4_final.pdf.

MOnarCH Project Consortium (2016). Project Deliverable D8.8.5 - Experiments on the MOnarCH Mixed Human-Robot Environment. http://users.isr.ist.utl.pt/∼jseq/MOnarCH/Deliverables/D8.8.5-v1c.pdf.

Naganathan, E. and Eugene, X. (2009). Software Stability model (SSM) for Building Reliable Real time Computing Systems. In *Procs. of the Third IEEE International Conference on Secure Software Integration and Reliability Improvement (SSIRI 2009)*. July 8-9.

Riek, L. (2012). Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. *Journal of Human-Robot Interaction*, 1(1):119–136.

Scassellati, B. (2005). Quantitative Metrics of Social Response for Autism Diagnosis. In *Procs of ROMAN'05*.

Sequeira, J. and Ferreira, I. (2014). The concept of [Robot] in Children and Teens: Some Guidelines to the Design of Social Robots. *International Journal of Signs and Semiotic Systems*, 3(2):35–47.

Ventura, R. and Ahmad, A. (2014). Towards Optimal Robot Navigation in Domestic Spaces. July 25 , Joao Pessoa, Brazil.

www.noldus.com (2016). Innovative solutions for behavioral research. Accessed March 2016.