Interactive spaces for children: gesture elicitation for controlling ground mini-robots

Patricia Pons[1]

ISSI Group, Departamento de Sistemas Informáticos y Computación (DSIC), Universitat Politècnica de València

Camino de Vera s/n, 46022, Valencia (Spain)

ppons@dsic.upv.es

Javier Jaen

ISSI Group, Departamento de Sistemas Informáticos y Computación (DSIC), Universitat Politècnica de València

Camino de Vera s/n, 46022, Valencia (Spain)

fjaen@upv.es

---

[1] Corresponding author

Interactive spaces for children: gesture elicitation for controlling ground mini-robots

Abstract.

Interactive spaces for education are emerging as a mechanism for fostering children's natural ways of learning by means of play and exploration in physical spaces. The advanced interactive modalities and devices for such environments need to be both motivating and intuitive for children. Among the wide variety of interactive mechanisms, robots have been a popular research topic in the context of educational tools due to their attractiveness for children. However, few studies have focused on how children would naturally interact and explore interactive environments with robots. While there is abundant research on full-body interaction and intuitive manipulation of robots by adults, no similar research has been done with children. This paper therefore describes a gesture elicitation study that identified the preferred gestures and body language communication used by children to control ground robots. The results of the elicitation study were used to define a gestural language that covers the different preferences of the gestures by age group and gender, with a good acceptance rate in the 6-12 age range. The study also revealed interactive spaces with robots using body gestures as motivating and promising scenarios for collaborative or remote learning activities.

1.   Introduction

Technology has been shown to be a promising and motivational tool for educational purposes with children. In this regard, over the last years there has been a shift from desktop-based interactions between children and digital services for learning, towards more ubiquitous and intuitive scenarios. This has favored the emergence of interactive spaces for education aimed at fostering children's natural ways of learning by means of play, exploration in physical spaces and tangible manipulation (Alborzi et al. 2000; Grønbæk et al. 2007; Marco et al. 2009; Mora-Guiard et al. 2016). For these reasons, these scenarios are not just purely virtual any more but instead are interactive spaces including advanced interactive modalities, such as augmented interactions, and non-digital artifacts, such as tangible devices of different kinds (Soler-Adillon et al. 2009; Bonarini et al. 2015; Garcia-Sanjuan et al. 2015). Among the wide variety of tangible interactive devices that are available in these spaces, children find robots attractive and intuitive. In fact, robots have been a popular research topic in the context of educational tools. Interactive activities between robots and children have the potential to foster physical activity, trigger creativity and provide a more engaging learning experience (Salter et al. 2006; Tanaka and Takahashi 2012; Wainer et al. 2014; Garcia-Sanjuan et al. 2015).

In order to provide users with more freedom and spontaneity in their interactions, advances in Natural User Interfaces (NUIs) have been applied in interactive spaces for adults (Wilson and Benko 2010; Benko et al. 2014; Jones et al. 2014; Fender et al. 2017; Vatavu 2017). Research on NUIs has widely covered how adult participants would like to interact with such novel interactive spaces. There is now a vast literature on gesture elicitation studies for adults in different contexts (Kray et al. 2010; Morris et al. 2010; Morris 2012; Vatavu 2012; Piumsomboon et al. 2013), which also demonstrates that user-defined gestures are usually better accepted by end-users (Morris et al. 2010; Pyryeskin et al. 2012; Dong et al. 2015). On the contrary, less research has focused on how children would prefer to interact in these scenarios. In the particular case of robots, research has also studied how adults would naturally interact using their full body for the navigation and control of these physical elements (Micire et al. 2009; Obaid et al. 2012, 2016; Cauchard et al. 2015). However, few platforms for children offer gestural and body language communication with these interactive elements. Gesture-based communication with robots would be especially relevant within interactive spaces, as for example, voice commands might not be accurate if there is ambient noise or there are many children playing at the same time. In addition, current scenarios in which children can interact using body language are usually based on virtual environments,

and the gestural languages have already been predefined by the developers. Therefore, studying how NUIs and body language communication could provide children with a more intuitive and motivational experience for controlling tangible robots within interactive scenarios would be a promising research line.

To the best of our knowledge, this paper presents the first gesture elicitation study with children regarding full body interactions to manipulate tangible ground robots. The outcomes of this elicitation study are manifold: first, the study identifies the preferred gestures and body language communication that children would like to use in interactive spaces with robots. It also outlines the variability of such gestures by age group and gender, proposing a gestural language that accommodates for the detected requirements. Secondly, the positive results from the questionnaires reveal that interactive scenarios with robots using body gestures are very promising and motivating for children. Finally, a set of future recommendations based on the elicitation study and the postquestionnaire results is outlined. This discussion aims to advance research in the design and development of promising interactive spaces for children based on natural interactions.

2.   Related work

This section firstly reviews related studies on children's communicative forms of interaction within interactive spaces and with robotic systems. Then, studies focused on gesture and full-body elicitation techniques with children are described, highlighting the need for more participatory design processes with children in these areas. Due to the limited number of gesture elicitation studies with children regarding interactive spaces and tangible robots, we propose revisiting the related work in this area focused on adult participants.

2.1.   Children's communication with robots and interactive spaces

Child Computer Interaction (CCI) (Read et al. 2008; Read and Markopoulos 2013; Hourcade 2015) research has expanded greatly over the last decades, partly due to the realization that children's mental models are different from those of adults. Hence, interactive technologies for children that aim to be successful need to adapt to these mental models and preferred ways of interaction. In this process of understanding and adapting technology to better suit children's requirements, cooperative inquiry and participatory design with children have been among the main pillars of the field (Druin 1999; Druin et al. 1999). These methods have been used in search of motivating and playful technology that encourages exploration and learning in children, and interactive spaces and robots have emerged as very promising tools for this purpose (Antle et al. 2009; Horn et al. 2011).

On one hand, interactive spaces are moving towards full-body interactions in which children can behave naturally using their whole body to interact with the system. This offers a more natural interactive modality while encouraging exercise and socialization. Back in 2000, the *KidsRoom* (Bobick et al. 1999) and *StoryRooms* (Alborzi et al. 2000) proposed interactive storytelling experiences for children that took place in an interactive real world room. Over the years, advances in technology have made these immersive scenarios more affordable and sophisticated. For example, interactive floors such as *iGameFloor* (Grønbæk et al. 2007) or *Lands of Fog* (Mora-Guiard et al. 2016) allowed full-body interactions for collaborative learning and socialization using computer vision techniques. Other interactive playgrounds for children that considered full body interactions have been the interactive slide (Soler-Adillon et al. 2009), gesture-based digital games (Pares et al. 2005; Gonzalez et al. 2009; Hsiao and Chen 2016) and hybrid digital and tangible games (Bonarini et al. 2015).

On the other hand, interactions between robots and children have been studied from different perspectives and offering several interactive modalities. However, few works have addressed embodied interaction techniques between children and robots. Those works in which embodied metaphors are present are usually focused on imitation and social interactions for therapy, e.g. socialization with robots for children with autism (Robins et al. 2008; Robins and Dautenhahn 2014; Wainer et al. 2014). Embodied interactions and body posture have also been analyzed to detect children's engagement and social interactions with robots (Michaud and Caron 2002; Salter et al. 2006; Sanghvi et al. 2011; Henkemans et al. 2017). However, all these studies consisted of interactive modalities already pre-programmed by the designers, and the children had to adopt them in order to interact with the robot. Some examples include verbal communication (Belpaeme et al. 2013), gesture imitation (Robins et al. 2008; Wainer et al. 2014), touch-based interactions (Michaud and Caron 2002; Salter et al. 2006; Robins and Dautenhahn 2014) or tangible-mediated interactions (Garcia-Sanjuan et al. 2015; Nacher et al. 2016).

The proposed interactive modalities have been quite varied as regards children's control and manipulation of a robot. *Topobo* robot (Raffle et al. 2004) required children to perform physical programming by example, i.e. assembling the robot and recording the desired movement by moving its physical parts, then playing the recorded movements. *LEGO sheets* proposed a visual programming language for the LEGO Programmable Brick, a physically assembled LEGO robot (Gindling et al.). This allowed the robot to be programmed with predefined sequences of actions, but did not provide real time control of it by the children. A similar proposal was *Robo-blocks* (Sipitakiat and Nusen 2012), which offered a tangible approach for programming a floor robot by assembling physical command blocks. However, it offered only four basic actions: forward, backwards, left and right. The *Tangibot* platform (Garcia-Sanjuan et al. 2015; Nacher et al. 2016) proposed a more collaborative real-time technological solution. It consisted of a tangible-mediated user interface for kindergarten children to manipulate a mobile robot. The work in (Tanaka and Takahashi 2012; Takahashi et al. 2012) also proposed a tangible interface for children in the form of a tricycle to remotely control a robot. The movements performed on the tricycle were mapped to movements of the tele-operated robot. However, it did have some technical limitations, as the child operating the robot could not perform fine-grained interactions and could not move further when he reached a wall, etc.

There is thus a shortage of studies on how full body interactions that work well in interactive spaces for children could be applied to manipulating robots in such spaces. We believe the study of robot control by gestural communication in an interactive space will offer promising opportunities to develop motivating and engaging technological scenarios. The following sections will review elicitation studies on both children and adults to give a general idea of the range of methodologies that could be applied for the purpose of this study.

2.2. Elicitation studies with children

Gesture elicitation studies with children have been used as a tool to include children in the design process of a system (Druin et al. 1999; Druin 2002; Guha et al. 2005) following a participatory design approach (Derboven et al. 2015). However, the relatively few studies in the area mostly focus on more traditional user input interfaces, such as tabletops (Rust et al. 2014) or digital videogames (Höysniemi et al. 2005).

Regarding interactive surfaces, a methodology proposed by Wobbrock et al. (2009) within adult elicitation studies has been adapted to study how children designed new gesture-based interactions on tabletops (Rust et al. 2014). In this study, after observing an animation representing some referent action, the children were asked to come up with a gesture to trigger that action.

Höysniemi et al. (2004, 2005) conducted an elicitation study in which children had to perform full body interactions to control a digital character in a virtual game. This work used a Wizard of Oz elicitation technique, showing this is a feasible methodology well accepted by children, despite the delay between the child's action and the reaction of the character.

With the emergence of depth-based tracking technology such as the Microsoft Kinect sensor, gesture elicitation studies with children have also evaluated how they would like to manipulate digital objects which appear on screen (Connell et al. 2013). Such systems allow for rapid prototyping and usually facilitate the gesture recognition task with built-in frameworks. However, they still have limitations in terms of open space tracking and gesture precision, which is a drawback for less restricted scenarios.

In the context of interactive environments for learning, bodystorming techniques have been used to elicit children's full body interactions to improve their communication with the system (Schaper et al. 2015; Malinverni et al. 2016). In these studies, children were asked to use their bodies to represent possible actions and relationships between elements in order to address issues related to air-pollution. Motivation was reported as high among the participants, although the proposed gestures were more focused on representing the different actors of the scenario rather than solving the design challenge. Although the authors acknowledge some of the representations cannot be directly mapped to the final system, full body interactions were a powerful tool within the participatory design process with children (Schaper and Pares 2016).

2.3. Elicitation studies with adults

In contrast to the reduced number of works regarding gesture usability studies with children, research focusing on gesture elicitation studies with adults has been extensive.

The widespread use of handheld devices motivated a great number of studies for eliciting natural gestures for mobile phones or tablets. Some of the works in the area include elicitation studies to use the back of the mobile device as an input (Shimon et al. 2015), or use motion gestures to trigger actions on the device (Ruiz et al. 2011). Considering more collaborative and interactive scenarios, several studies have addressed how users would like to interact with different devices in a multi-display environment in order to transfer files or share content (Kray et al. 2010; Kurdyukova et al. 2012; Seyed et al. 2012). The methodologies of these elicitation studies were in the form of interviews or guided tasks: the participant was asked to perform a gesture after reading a description of the task or observing a short image/video of the results for the proposed action.

Hand gestures for vertical displays (TVs or interactive walls) have also been elicited using different approaches. For example, online surveys in which participants propose gestures for a specific command using open-ended answers (Dong et al. 2015). Another common procedure consists of the participant performing a gesture after observing a short video demonstration of the proposed command (Vatavu 2012). And finally, Wizard of Oz methodologies, in which the referent action is initially shown to the user and also is triggered while he performs the action (Morris 2012; Lee et al. 2013). With the emergence of tabletops, gesture elicitation studies were also conducted to elicit user-defined gestures for interactive surfaces (Epps et al. 2006; Wobbrock et al. 2009). Methodologies reported in the literature using these platforms consisted of showing an animation of a referent action to the participant, and asking him to perform a gesture that would prompt that action.

Moving to more immersive interactive spaces, elicitation studies for collocated interactions with wearables have also been conducted (Jokela et al. 2016). Regarding digital games, full-body gesture elicitation procedures (Kistler and André 2013) required the participant to play a digital game standing in front of a screen. Whenever the game script stops and displays different options for the character to continue the story, the participant is asked to perform a gesture that represents each option. User defined gestures have also been elicited in Augmented Reality (AR) scenarios. Using a head-mounted device, participants are shown an animation of the desired result and then asked to perform a suitable gesture to cause that effect (Piumsomboon et al. 2013).

Several elicitation studies focused on adult manipulation of either humanoid robots (Obaid et al. 2012, 2014) as well as flying drones (Cauchard et al. 2015; Obaid et al. 2016). The methodologies adopted in these studies usually follow a Wizard of Oz approach: (1) the user is shown a video of the robot/drone performing a referent action or reads its description on a card, (2) the user performs a gesture that represents the referent action, and (3) the robot/drone performs the desired action. Obaid et al. (2012) also proposed a gesture taxonomy for the classification of the elicited gestures, which has been influenced by the taxonomies used by Wobbrock et al. (2009) and Karam and Schraefel (2005) in previous gesture elicitation studies. In addition, there is an overwhelming amount of studies evaluating drone/robot manipulation techniques for adults, e.g. hand gestures (Sato et al. 2007; Konda et al. 2012; Nahapetyan and Khachumov 2015), full-body interactions (Pfeil et al. 2013; Sanna et al. 2013), wearables (Sugiyama et al. 2011), gaze (Hansen et al. 2014), speech (Stiefelhagen et al. 2004; Fernandez et al. 2016), etc. This contrasts with the few works on the topic with children as participants. For these reasons, and considering how well Wizard of Oz elicitation techniques work for the user in several contexts, we propose a Wizard of Oz elicitation study to evaluate how children would like to manipulate a robot using full-body interactions.

3. Elicitation Study

The main goal of this study is to identify user-defined gestures to control and move a non-humanoid ground mini-robot. This section will describe the study, procedure and the results obtained. The research questions of this work may be formulated as follows:

- RQ1: Is the cognitive load of gestural communicative interaction with robots affected by gender?
- RQ2: Is the cognitive load of gestural communicative interaction with robots affected by group age?
- RQ3: Are there any differences in the kind of gestures elicited by gender and age group?
- RQ4: Is there any gestural language that provides agreement across gender and age?
- RQ5: Is there any difference in the perceived enjoyment of the robot by gender and age group?
- RQ6: Is there any difference in the perceived easiness of the elicitation task by gender and age group?
- RQ7: Is there any difference in the motivation to play by gender and age group?
- RQ8: Is gestural control of robots a promising tool for remote playful activities?

- RQ9: Is gestural control of robots a promising tool for collaborative playful activities?

Section 4 will review these research questions reflecting on the results of the proposed study, with the aim of providing a useful discussion that helps future researchers in the field to design and develop suitable interactive spaces with robots for children.

3.1. Participants

The participants were 61 children from a public primary school (*Col·legi Públic Vicente Gaos*), and the sessions were conducted during school teaching hours. The participants were classified into three different age groups: early, medium and advanced stage of primary education. The first group (G1) were 21 children (14 boys and 7 girls) aged 6-7 years old (Mean (M) = 6.24, Standard Deviation (SD) = 0.436). The second group (G2) were 20 children (10 boys and 10 girls) aged 9-10 years old (Mean (M) = 9.25, Standard Deviation (SD) = 0.444). The third group (G3) were 20 children (10 boys and 10 girls) aged 11-12 years old (Mean (M) = 11.45, Standard Deviation (SD) = 0.605). This division aligns with children developmental theories, allowing to observe differences in the three stages proposed by Piaget (1973). Children in the age group between 4-7 years old are in the preoperational period, in which they primarily learn thorough imitation and play. Then the concrete operational stage follows, between 7-11 years old, and is characterized by the appropriate use of logic. Finally, the formal operational stage starts at around 11 years old. In this stage, individuals move beyond concrete experiences and begin to think abstractly.

Before starting the elicitation study, the children were asked whether they had previously used the Microsoft Kinect sensor when playing videogames. Most of them did not know the device before the session, however several children reported having played games with this or a similar sensor: 5 children in G1, 2 children in G2 and 10 children in G3. The children were also asked whether they had previous experience with robots or drones of any kind. In this regard, 4 children in G1, 3 children in G2 and 5 children in G3 reported either owning or having played with a robot/drone before. However, the interaction modalities with such devices had been based on controlling the device with a smartphone/tablet, but never with NUIs such as body tracking.

3.2. Set-up

The study was conducted following a Wizard-of-Oz approach. A Microsoft Kinect v2 sensor was placed on a table, facing a play area of 200cm wide x 300cm long in which the activities took place. The device video-recorded the activity of the participants for subsequent analysis. The robot used in the study was a Parrot Jumping Sumo mini-drone, placed initially in front of the Kinect. The children were required to stand behind the robot facing the Kinect sensor to start the session, but they could move freely within the play area during the rest of the activity. The researcher stood behind the play area manipulating a smartphone to control the robot (**Fig. 1**).
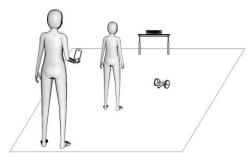

Fig. 1. Sketch of the Wizard-of-Oz set-up.

3.3. Procedure

Each child was asked to perform three tasks, with no previous training for any of them, in order to elicit the gestures that children would naturally want to perform to control the robot in each scenario. Task 1 consisted of eliciting gestures for specific and discrete movements of the robot. For this purpose, a set of 11 common actions for controlling mini-robots was randomly presented to each child (using a balanced Latin Square) to avoid order effects in and between groups. Table 1 shows the 11 actions and a brief description of each one. Each child was asked to come up with a gesture for each of the actions of the robot described in Table 1. For each action, Wobbrock's procedure was followed: (1) the robot performs

a specific movement while the child observes, i.e. the referent, (2) the child is asked to perform any gesture or movement she wants so that the robot performs the referent action, i.e. the sign, (3) once the child starts to perform the gesture, the wizard (researcher) starts moving the robot according to the specific referent while following the child's gesture.

Table 1. List of actions (referents) the mini-drone can perform.

Action of the robot (referent)	Code	Description
Move forward	F	The drone advances forward
Move backwards	B	The drone advances backwards
Turn left	L	The drone turns left on-site
Turn right	R	The drone turns right on-site
Move left	ML	The drone turns left while advancing
Move right	MR	The drone turns right while advancing
Speed up	SU	The drone speeds up while it advances forward
Speed down	SD	The drone slows down while it advances forward
Jump	J	The drone performs a jump
Stop	S	The drone stops moving
Spin	SP	The drone turns around several times

In contrast to Task 1, Task 2 and Task 3 required the children to perform a series of gestures over time in order to move the robot continuously following a predefined path on the ground marked with colored cardboards (42cm x 59.6 cm). All children were asked to perform both tasks, alternating the order between the participants to avoid ordering effects. The path in each task was designed to study the different modalities in which a child could give orders to a robot depending on the type of movement she wanted to perform. In Task 2, the path was formed by three straight segments, including one left turn and one right turn (see Fig. 2).

Fig. 2. Path from Task 2.

Task 3 consisted of an irregular path, including one left turn, one right turn and three special cardboards (see **Fig. 3**). On the blue cardboard, the children had to make the robot spin (SP) or jump (J) and on the orange cardboards, they had to make the robot speed up (SU).

During Tasks 2 and 3, the children could perform gestures to give the robot any signs they wanted, for example, they could stop it and prevent it from leaving the path, or correct its trajectory if they thought that the robot might leave the path.

Fig. 3. Path from Task 3.

After the three tasks, the children were asked to fill in a brief questionnaire (Read et al. 2002) on their opinion of the activity, the control of the robot and potential improvements (see Table 2).

Table 2. Postquestionnaire and answer options.

Code	Question	Answer options
Q1	How much did you enjoy the game?	5-point Likert scale
Q2	How easy was it to control the robot?	5-point Likert scale
Q3	Would you like to play with the robot in school again?	Yes / No / Maybe
Q4	Would you like to play with the robot out of school again?	Yes / No / Maybe
Q5	How would you like to play again?	Robot in the same room / Robot in a different room
Q6	Whom would you like to play with again?	Alone / Friends / Family / Other
Q7	What would you change in the game to make it better?	Open answer

3.4. Results

3.4.1. Time

Table 3 and **Fig. 4** sum up the average time the participants spent on each task by age group and gender. Start and end times for each task were computed from the time the researcher told the child that he had control of the robot to the time when the robot stopped moving after the child's last sign. The measured time does not intend to evaluate how fast children performed a gesture, in the sense of evaluating their performance in a usability context. Instead, this time measure aims to evaluate the cognitive effort in learning how to use the robot. This cognitive load would comprise the children's understanding, planning and execution of the gesture, hence the proposed time measurement would be more aligned with assessing the spontaneity of the children's interactions and learning curve.

We used a Mann-Whitney U test analysis to compare differences between the independent groups of participants regarding the time they spent in each task (Mann and Whitney 1947). In the youngest age group (G1), the female participants spent on average more time than the male participants on each task. However, no statistically significant differences were found between genders in G1 according to the Mann-Whitney U test (T1: $U = 41.000$, $p = 0.550$; T2: $U = 38.000$, $p = 0.411$; T3: $U = 25.000$, $p = 0.073$). In G2, the girls spent on average more time than boys only on Task 1, however there were no significant differences according to the Mann-Whitney U tests ($U = 31.000$, $p = 0.151$). Male participants spent more time performing the activity in Tasks 2 and 3. No significant differences were found in Task 2 ($U = 43.500$, $p = 0.622$), but Mann-Whitney U tests showed statistically significant differences by gender in Task 3 ($U = 23.5000$, $p = 0.044$). In this case, the Box-and-Whisker plots for this task show two outliers in the male participant group, corresponding to two boys who spent considerably more time on Task 3 than the rest of their colleagues. In G3, the female participants spent more time on average on all tasks. While no statistically significant differences were found for Tasks 1 and 3 (T1: $U = 41.000$, $p = 0.496$;

T3: U = 49.500, p = 0.9), they were found between female and male participants in G3 for Task 2 (U = 22.5000, p = 0.037). Although no outliers were identified in the Box-and-Whisker plots, this difference could be explained by analyzing the performance of the two female participants who took more time at this task. In both cases, the researcher performing the Wizard-of-Oz approach had difficulties understanding the intended meaning of the children's gestures. Nevertheless, considering the three age groups together, no statistically significant differences were found between the male and female participants in any task, according to Mann-Whitney U tests (T1: U = 415.500, p = 0.528; T2: U = 343.500, p = 0.093; T3: U = 440.500, p = 0.788)

Table 3. Average time (seconds) for each task by age group and gender.

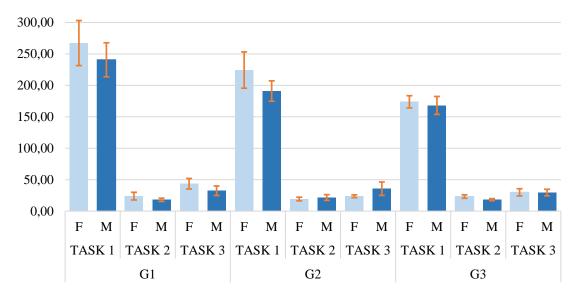|		Task 1	Task 2	Task 3
G1	F	$\bar{x} = 267.29, \sigma = 71.679$	$\bar{x} = 23.86, \sigma = 12.240$	$\bar{x} = 43.57, \sigma = 16.772$
	M	$\bar{x} = 240.71, \sigma = 54.109$	$\bar{x} = 18.00, \sigma = 5.144$	$\bar{x} = 32.43, \sigma = 15.210$
G2	F	$\bar{x} = 224.40, \sigma = 57.707$	$\bar{x} = 19.30, \sigma = 5.774$	$\bar{x} = 23.60, \sigma = 4.477$
	M	$\bar{x} = 191.00, \sigma = 32.435$	$\bar{x} = 21.70, \sigma = 9.141$	$\bar{x} = 35.70, \sigma = 21.308$
G3	F	$\bar{x} = 173.90, \sigma = 19.393$	$\bar{x} = 23.20, \sigma = 5,554$	$\bar{x} = 29.90, \sigma = 11.406$
	M	$\bar{x} = 168.10, \sigma = 28.544$	$\bar{x} = 18.30, \sigma = 3.234$	$\bar{x} = 29.60, \sigma = 10.426$

Fig. 4. Average time (seconds) for each task by age and group gender.
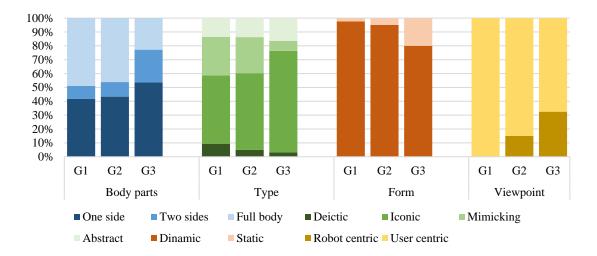
Mann-Whitney U tests were also performed to detect statistically significant differences between age groups in each task. There were no differences between age groups regarding the average time spent on Task 2 and Task 3. Statistically significant differences were found only in Task 1 between all groups: G1 and G2 (U = 122.5, p = 0.022), G1 and G3 (U = 30.500, p = 0.000), and G2 and G3 (U = 89.000, p = 0.003).

3.4.2. Gesture taxonomy

This section describes the taxonomy of gestures elicited from the study and the distribution of the observed gestures in the taxonomy by age group and gender. Table 4 sums up the gesture taxonomy derived from this study, which has been influenced by the taxonomies used by Wobbrock et al. (2009), Karam and Schraefel (2005) and Obaid et al. (2012). We manually classified all gestures in Task 1 according to the dimensions *body parts*, *type* and *form*, while Task 2 and Task 3 were used to classify the gestures according to the *viewpoint* dimension.
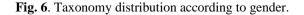
Table 4. Gesture taxonomy.

Body parts	One side	The gesture is performed with one hand, one arm or one leg
	Two sides	The gesture is performed with both hands or both arms
	Full-body	The gesture is performed using the whole body, or a combination of hand/arm and leg movements

Type	Deictic	The children indicates the movement the robot has to follow, either by pointing at its final destination or by performing a dynamic movement indicating the path/movement to be followed
	Iconic	The gesture visually depicts the movement to be performed or is a visual metaphor of it
	Mimicking	The children imitates the movement of the robot
	Abstract	The gesture has no direct mapping with the robot's movement
Form	Static	After a short preparation phase in which the user prepares the desired gesture, this gesture does not entail movement
	Dynamic	The gesture entails movement
Viewpoint	User-centric	The gesture is performed from the user's point of view (egocentric)
	Robot-centric	The gesture is performed from the robot's point of view (allocentric)

Fig. 5. Taxonomy distribution according to age groups.

Fig. 6. Taxonomy distribution according to gender.

The children were free to perform any gesture or movement they wanted to for each action of the robot in Task 1. Hence, they could move anything from a hand to their whole body, e.g. walking besides the robot. Their elicited gestures have been analyzed regarding the body parts used to perform each action in Table 1 according to the taxonomy in Table 4. The three different categories in which gestures can be classified attending to the body parts involved are: one side, two sides or full body interaction. As can be observed in **Fig. 5**, most of the children used either one-side gestures or full body interactions, and this happened across all the different actions of the robot (see **Fig. 7**). Gestures involving both sides of the body, such as using either both hands or both arms, were less frequent. According to the Pearson Chi-Square tests performed on the 662 gestures classified in Task 1, there is a statistically significant association between

the participant group (children's age) and the preferred body parts they used to perform the gestures (χ (4) = 46.222, p = < .001). Post hoc analysis with adjusted Bonferroni correction showed that full body interactions were less frequent in children in G3 than in younger children in G1 (χ (2) = 38.001, p = < .001) and G2 (χ (2) = 30.920, p = < .001). In addition, children in G3 used both sides of their body statistically more frequently than children in G1 and G2 (p = < .001). There was also a statistically significant association between gender and the preferred body parts (χ (2) = 10.265, p = .006). In this regard, **Fig. 6** shows that girls tended to use their whole body more frequently to interact (43% of girls against 36% of boys), while boys tended to use just one side of their body (52% of the boys against 40% of the girls).
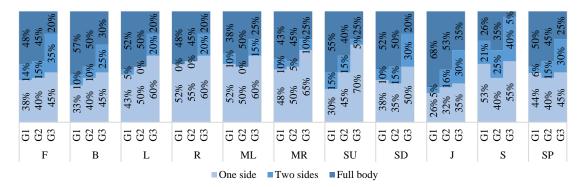

Fig. 7. Distribution on the use of body parts for different actions by age group.

Another relevant aspect to consider within the elicitation study was the type of metaphors or abstractions the participants used in their gestures to give meaning to a specific action (or referent). This was somehow a new definition of a shared non-verbal language between the participant and the robot. One of the main goals of this study was to identify the type of gestures children would feel more natural to use when "*speaking*" to a robot by body language. The gestures were classified into four different types (see Table 4): deictic, iconic, mimicking and abstract. *Deictic* gestures are those in which children point with their fingers to a specific location or direction and may involve movement. In this case, the children point and trace the path the robot has to follow to perform a specific action. *Iconic* gestures are metaphoric gestures that imply either a direction, speed or visual depiction of the movement: moving the hand rapidly, moving the hand/arm from left to right, etc. *Mimicking* gestures are those in which children perform the same movement as the robot, as if they were showing the robot how to do it. *Abstract* gestures have no direct mapping with the direction, speed or movement of the action.

According to the Pearson Chi-Square tests performed on the data, there was a statistically significant association between the participant's age group and the type of gesture the children performed (χ (6) = 48.014, p = < .001). Post hoc analysis with adjusted Bonferroni correction showed statistically significant differences between G1 and G3 (χ (3) = 44.253, p = < .001), and between G2 and G3 (χ (3) = 30.133, p = < .001). Children in G3 usually performed more iconic gestures than children in G1 and G2 (p = < .001). Mimicking gestures were rarely used by children in G3 (p = < .001), while children in G1 and G2 made similar use of them. There was also a statistically significant difference between female and male participants regarding deictic gestures (χ (3) = 15.506, p = .001). Girls tended to use deictic gestures statistically less than boys did (p = .002), as can be observed in **Fig. 6**. **Fig. 8** also shows the distribution of gesture types for the different actions and age groups.
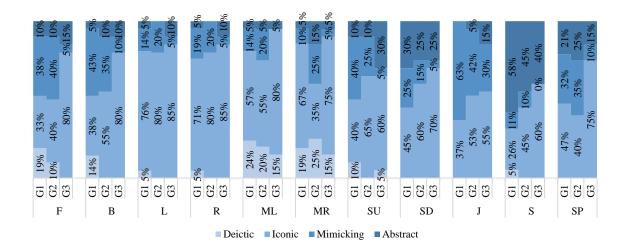
Fig. 8. Distribution of types of gestures by action and age group.

The form of the gestures could be static or dynamic. Static gestures do not entail movement after an initial preparation phase in which the child places himself in the desired posture/gesture, e.g. horizontal still arm towards the front to indicate *forward* movement, as in **Fig. 9**. Instead, dynamic gestures accompany the posture with movement, e.g. moving arm and hand from right to left to indicate movement in that direction, as in **Fig. 10**. **Fig. 5** shows that children barely used static gestures and that all three age groups preferred dynamic gestures. However, the children in G3 used static gestures slightly more than the younger groups, and post hoc analysis with adjusted Bonferroni correction showed statistically significant differences between G1 and G3 (χ (1) = 6.509, p = 0.011) in this regard. Girls used dynamic gestures (94.44%) more frequently than boys (88.24%) on average. However, there was no statistically significant difference between genders, according to Pearson Chi-Square tests (χ (1) = 1.415, p = 0.234).

The viewpoint dimension was a very interesting aspect to evaluate. According to the taxonomy, a robot-centric gesture is one in which the sign is given according to the robot's coordinate system, e.g., the robot has to move towards the left side of the room and it is already facing the left wall, hence the user performs a *Forward* gesture such as in **Fig. 9**. On the other hand, user-centric actions are given according to the user's coordinate system, regardless of the orientation of the robot, as in **Fig. 10**. The referent actions provided in Task 1 started always with the robot in the same position and orientation, and it was difficult to assess whether children were considering a user-centric or a robot-centric point of view, as both coordinate systems mapped to each other in this initial configuration. In order to assess children's preferences towards robot-centric or user-centric gestures, we manually classified this dimension in Task 2 and Task 3 for each child (122 samples). The nature of tasks 2 and 3 was different from the elicitation activity in Task 1 as the referents in Task 2 and 3 were not discrete actions of the robot. Instead, in Tasks 2 and 3 the children had to move the robot over a predefined path on the ground. This allowed us to observe clearly which coordinate system they adopted as reference for their interactions, as well as whether they changed the gestures they were performing in Task 1 to new ones better suited to the nature of these new tasks.

Overall, the children usually adopted a user-centric view, with an increase in the adoption of a robot-centric view with age. According to the Pearson Chi-Square tests performed on the data, there was a statistically significant association between the participant's age group and the viewpoint the children adopted (χ (2) = 16.473, p < .001). Post hoc analysis with adjusted Bonferroni correction showed statistically significant differences between G1 and G2 (χ (1) = 6.797, p = 0.009) and between G1 and G3 (χ (1) = 16.222, p < .001). This confirms that the observations in **Fig. 5** are statistically significant: no children in G1 used a robot-centric view, while 15% of the children in G2 and 33% of children in G3 preferred this viewpoint instead of a user-centric one. In relation to gender, on average girls adopted a user-centric view (96.30%) more frequently than boys did (75.00%). In this regard, Pearson Chi-Square tests confirmed a statistically significant difference between female and male participants regarding the adopted viewpoint (χ (1) = 10.382, p = 0.001).

Children were free to use the gestures they wanted for Tasks 2 and 3, even if those gestures were different from the ones they had elicited during Task 1. Children usually performed the same or very similar gestures between Task 1 and Task 2, and between Task 1 and Task 3, with only 34.43% of children changing most of the gestures they had defined in Task 1 in these two new activities. Some children

interchanged the gestures they used for L/R with the ones they used for ML/MR (7 children), and a less common modification was to interchange the F and SU gestures (6 children). The most common pattern was to change the gestures they had used in Task 1 for a deictic or mimicking approach: instead of using a gesture to indicate each separate action to the robot to follow the path, they perceived the path as a whole action (13 children). In this way, instead of giving a separate action for each section of the path, children pointed the path to the robot as a continuous action, or walked over the path/near the robot to define the directions it should follow. However, once they had established these new gestures for either Task 2 or Task 3, children kept those gestures the same between Tasks 2 and 3, with only 5 children again the set of gestures between tasks.

Fig. 9. Example of robot-centric gesture.

Fig. 10. Example of user-centric gesture.

3.4.3. Gesture language

This section describes the gestural language elicited from the children's gestural data collected during Task 1 of the study for controlling a ground mini-robot. The selection of a suitable gesture for each control action took into account the agreement levels among the children's defined gestures. The process of selecting a suitable gesture for a control action is as follows:

- For each control action a we identify a set P_a that contains all the proposed gestures.
- The proposed gestures in P_a are then grouped into subsets of conceptually identical gestures $P_{i1...N}$, where i is a subset that contains conceptually identical gestures and N is the total number of identified subsets. A conceptually identical gesture can be either the same identical gesture, or a gesture that entails the same movement and meaning, e.g. *moving left hand from left to right* would be conceptually identical to *moving left hand with extended left arm from left to right*.
- The representative gesture for the control action a is identified by selecting the largest subset P_i.

To further evaluate the degree of agreement among the participants regarding the selected user-defined sets, we followed the process proposed by Vatavu and Wobbrock (2015, 2016) and used AGATe[2] software to compute an agreement score for each control action.

Table 5 reports the agreement levels for each action by participant group. The gestures with the highest agreement rates among all the participants were Jump (J), Stop (S), Spin (SP), Forward (F) and Backward (B). On the other hand, agreement rates for gestures involving turns (L, R, ML, MR) and speed changes (SU, SD) were quite low. However, it was observed that for those actions with low agreement rates, there were two or three gestures with similar frequencies instead of a single preferred gesture. For example, for ML and MR actions, around 15-23% of the time the children chose to move their hand or arm from left to right (or right to left), while 16-20% of the time they chose to draw a curve in the air. By considering the two or three most frequent gestures for each action, we were able to provide a consistent and robust

---

[2]http://depts.washington.edu/madlab/proj/dollar/agate.html

gestural language that conforms to almost 50% of the participants in the three age groups. The most frequent gestures for each action are illustrated in Fig. 11 with their respective frequencies.

3.4.4. Gesture analysis

As can be observed, the proposed gestural language has two/three gestures for each action that usually fall into a mimicking or iconic approach. Therefore, it could be argued that two different gestural languages could be proposed: one based on a mimicking approach and another one for an iconic/abstract approach. However, the observed results from the participants did not show any clear pattern in this regard. While several children used a majority of either mimicking or iconic gestures, this was not a common observation, especially in G1 and G2. These two groups with younger children presented a great variability in the type of gestures each participant preferred to perform for each action. The participants in G3 usually showed a clear preference towards iconic gestures, however not even half of them used iconic gestures for most of the proposed actions. The most frequent behavior observed across genders and age groups was that children varied their gestures from one approach to another. Hence, the proposed gestural language aims to accommodate to this variability for a wide age-range population.

The comparison of the elicited gestural language presented in Fig. 11 with previous elicitation studies regarding adults and children provides many interesting observations. For example, related elicitation studies with children to control digital objects using full-body interactions also highlight pointing (deictic) and walking (mimicking) movements as common preferred gestures (Connell et al. 2013). Nevertheless, elicitation studies with children usually detect great variability due to both individual preferences as well as to variations in the same type of gesture, as observed also in Connell et al. (2013). For example, a deictic gesture might be performed by different children in different ways: using one-finger pointing, using the palm of the hand or even moving the whole arm.

Several differences can be found when comparing the described gestural language in this manuscript with elicitation studies with adults for controlling robots. The observed differences could also be affected by the type of robot used in the study and the movements it could perform. In this regard, there have been elicitation studies with adults regarding full-body interactions for controlling a humanoid robot (Obaid et al. 2014), as well as a flying drone (Obaid et al. 2016). Regarding the gestures for controlling a flying drone, adults usually preferred two-hand interactions, which was the least preferred body movement performed by children (Obaid et al. 2016). In addition, adults barely performed mimicking gestures to control a flying drone, while this was one of the preferred type of gestures for children as can be observed in Fig. 11. In the case of humanoid robots, more mimicking and iconic gestures were observed but adults overall favored deictic gestures, which is the least preferred type of gesture for children. In this context, the gestures for the F, B, S, L and R actions seem quite similar in both studies, but overall children displayed a more varied set of gestures, with less preference towards two-hand gestures.
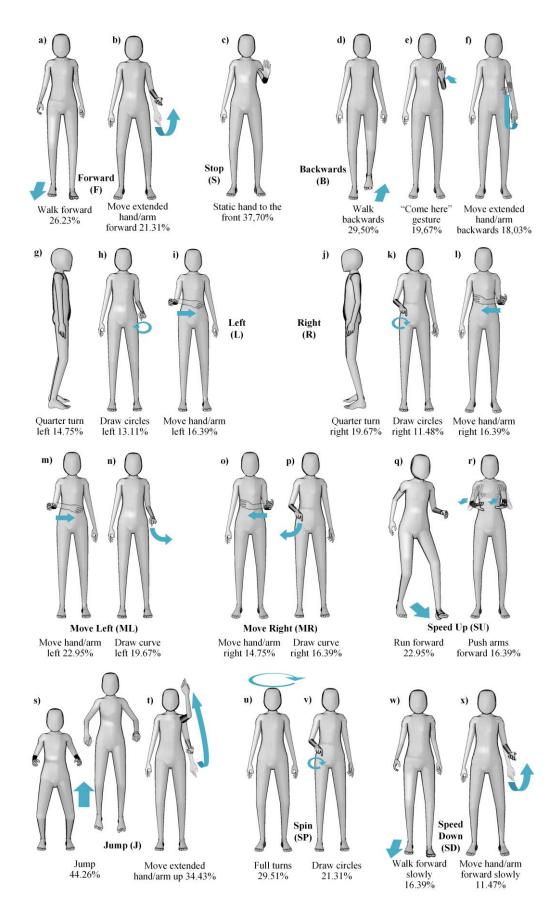
Fig. 11. Gestural language with most frequent gestures for each action and occurrence frequencies.
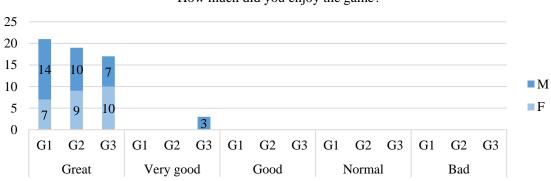
Table 5. Agreement level of gestures for the different actions proposed in T1 by participant group.

Action	G1	G2	G3	All groups
F	0.176	0.142	0.111	0.125
B	0.205	0.179	0.153	0.155
L	0.057	0.089	0.042	0.070
R	0.052	0.147	0.042	0.081
ML	0.138	0.053	0.089	0.098
MR	0.067	0.053	0.042	0.061
SU	0.186	0.100	0.084	0.104
SD	0.076	0.053	0.037	0.048
J	0.390	0.337	0.195	0.309
S	0.071	0.184	0.263	0.161
SP	0.090	0.189	0.147	0.133
Average	0.137	0.139	0.110	0.122

3.4.5. Postquestionnaire results

After the three tasks, children were asked to complete a postquestionnaire about their experience and thoughts. The questions are reported in Table 2, and this section outlines the main results.

Regarding question Q1, all the children in G1 and G2 and almost all those in G3 reported that the activity with the robot was "Great". Only three children in G3 did not select the highest option in the Likert scale, however they also reported a very positive experience. Overall, the activity seemed to be very amusing and enjoyable for the children (see **Fig. 12**), with no statistically significant differences between age groups nor gender.

Fig. 12. Results for Q1 in the postquestionnaire by participant group and gender.

The children were also asked about the perceived easiness of controlling the robot (Q2). Most of them reported that it was extremely easy (see **Fig. 13**), especially those in G1 (aged 6-7 years old). None of them found it difficult and only one child from G2 ranked this task as having a normal difficulty level. The Mann-Whitney U test only showed statistical significant differences by age groups between G1 and G2 (U = 113.5, p = 0.002) as well as between G1 and G3 (U = 103.0, p = 0.000).
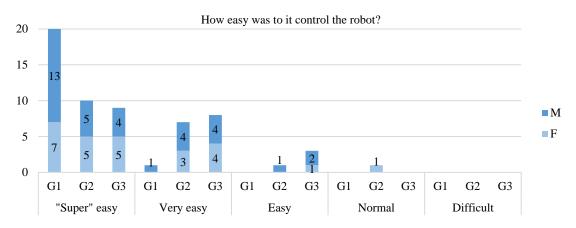
Fig. 13. Results for Q2 in the postquestionnaire by participant group and gender.

Another important aspect to consider was whether the children would like to play with the robot again and in which contexts. **Fig. 14** and **Fig. 15** show that they were very enthusiastic about playing with the robot again either in or outside the school, without significant differences. In addition, the robot used in the experiment had a camera that allows the user to observe on a screen what the robot is seeing in real time. This supports remote control of the robot: the user could give commands to the robot without having visual contact with it, but could guide it thanks to the built-in camera. This interactive scenario would be really useful for situations in which children cannot leave their physical location, as in the case of long-term hospitalized children. Hence, Q5 was aimed towards exploring the acceptance of children towards this kind of remote activity in order to design such a playful experience. The two interactive modalities with the robot, i.e. on-site vs. remote, were explained to all children in the questionnaire in order to ask them whether they would like to play with the robot again in the same room, control it remotely from a different room, or both (Q5). **Fig. 16** and **Fig. 17** show that most of the children were curious about exploring this kind of remote interaction, mentioning that in this way they could spy on their friends while they are in classroom, or they could use the robot to move things around and bring them to another person. The children in all the groups either preferred to play again only with the remote version, or marked both modalities, which opens a promising path towards the development of remote interactive systems for children based on NUIs.
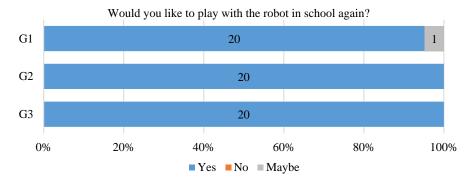

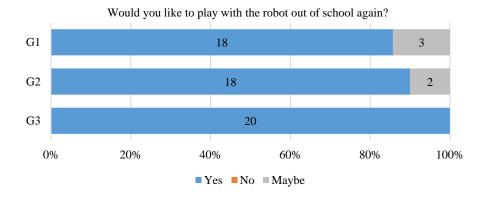**Fig. 14**. Results for Q3 in the postquestionnaire by participant group.

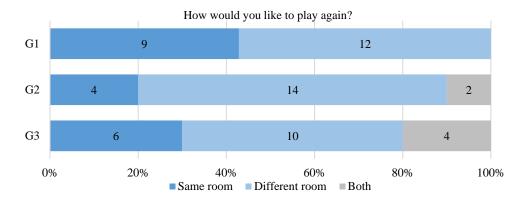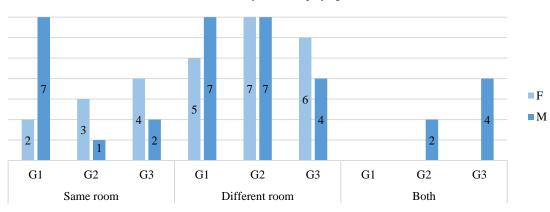**Fig. 15**. Results for Q4 in the postquestionnaire by participant group.

Fig. 16. Results for Q5 in the postquestionnaire by participant group.

Fig. 17. Results for Q5 in the postquestionnaire by gender and participant group.

Another promising opportunity for gesture-based control of mini-robots would be to create collaborative activities that foster educational values such as collaboration, communication, trust, etc. For these reasons, the children were asked whether they would like to play alone or with other people, such as friends, family or others (Q6). They were able to choose several options in this question. **Fig. 18** shows the histogram of each answer per participant group and gender. As can be observed, the shape of the distributions by age group and by gender are similar. Most of the children reported they would like to play with their friends, family and even others (instructor, teacher, etc.). The ones who preferred to play alone were mostly boys and only two girls preferred this option. However, only six out of these 10 children chose the solo option only, while the remaining four indicated they would also like to play with their friends or family.
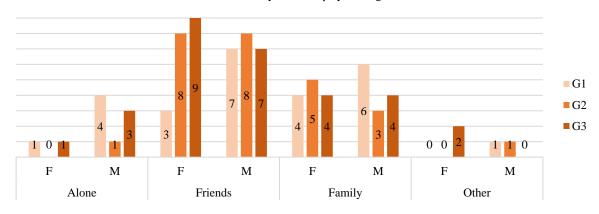
18

Whom would you like to play with again?

Fig. 18. Results for Q6 in the postquestionnaire by participant group and gender.

At the end of the session, children could also propose any changes that they would like to incorporate to the game or the robot (Q7). Several children reported that they would love to have several robots and be able to play with their friends. Some children commented they would like the robot to have more actions, and others asked for bigger or more complicated paths as well as incorporating obstacles.

4.   Discussion and Future Recommendations

The study explored the comprehensive set of research questions defined in Section 3 with the aim of advancing research and improve the development of future gestural languages for children depending that are suitable to their capabilities and preferences. This section will review the defined research questions in Section 3, providing recommendations for future designs.

4.1.   Evaluation of age and gender in gestural interactions

The answer to RQ1 (whether there are time differences by gender) is negative. Although there were differences in G1 and G3, they were explained by either communicative issues in the Wizard-of-Oz approach or outlier participants. Moreover, considering the three age groups together, the time performance using gestural communicative interactions with robots was similar for all the participants, regardless of gender. This is a promising result as it allows us to define a gestural language that could be used by a large population.

Regarding RQ2 in Section 3.4.1, significantly statistical differences regarding time were found in Task 1 between all age groups. However, these results are still positive as no differences were found in Tasks 2 and 3, which took place later. It seems that all the participant groups required different lengths of time to become familiar with the robot and the gesture signs they were defining, which can be explained by differences in cognitive skills relative to their age (Piaget 1973). However, once they got used to the device after Task 1, all the groups exhibited a comparable time spent in Tasks 2 and 3, which involved controlling the robot and giving it signs to move over a predefined path. In this regard, finding a suitable gestural language would mean that children of different ages in the 6-12 age range could perform equally well if provided with the appropriate communicative metaphors, in spite of their different cognitive and motor skills.

The elicited gestures were analyzed according to the taxonomy in Table 4 in order to answer RQ3 (whether age and gender influenced the kind of gestures according to the taxonomy). In this regard, several differences were found. For example, younger children preferred mimicking gestures while older children tended to produce iconic gestures instead. This is also related with the younger children's preferences for full-body gestures, because these full-body gestures usually mimic the robot's movements. In addition, the form and viewpoint of the elicited gestures was also affected by age. Overall, the children clearly preferred dynamic gestures and a user-centric viewpoint. However, as they get older there is an increase in the percentage of static gestures and they start adopting a robot-centric viewpoint more frequently. This latter effect has also been observed in other studies with children (Piaget 1956; Connell et al. 2013). Regarding gender differences, girls used fewer deictic gestures than boys, but adopted a user-centric viewpoint more frequently than their male counterparts. This could be due to gender differences in the abstraction of space and previous exposure to technological activities that entail

spatial representations, such as playing videogames or controlling a robot (Subrahmanyam and Greenfield 1994; Voyer et al. 1995). In addition, boys preferred to use just one side of their body to interact, while girls tended to use their whole body.

Despite the differences by age and gender, the set of gestures in **Fig. 11** that emerged from the elicitation study can allow for all the aforementioned differences. In this respect, the answer to RQ4 is affirmative, as we proposed a gestural language that provides agreement across gender and age. Implementing the two or three most common gestures for each of the proposed actions would allow to account for the different communicative modalities found. For example, either mimicking or iconic movements could be performed for most actions. This could support progressive change from mimicking gestures to iconic ones due to age differences. Regarding body parts, we could implement each of the iconic gestures in our gestural language with two versions: one hand/arm or two hands/arms. In this way, each action would have a full-body gesture as well as gestures that can be performed with one or two sides of the body. This would cover a wide range of preferences depending on age and gender, as shown in RQ3. Considering the form, less than 20% of the elicited gestures were static, and most of those happened in the preferred gesture for the Stop (S) action, which has been included in the gestural language. Therefore, most of the gestures in the proposed language would be dynamic. The speed of the movement would also be considered in order to allow for actions such as SD, which consists of similar, but slower, gestures as F. Perhaps the most complicated decision would be whether to use a user-centric or a robot-centric viewpoint. Our findings are consistent with the results of (Piaget 1956; Connell et al. 2013), which described how children gradually change from egocentric to allocentric gestures over time. In the age range considered in this study, the majority of participants used a user-centric viewpoint even in the oldest age group. However, if we also want to adapt to the robot-centric viewpoint preference, a possible solution is to have an initial "analysis" phase before the real interaction starts. This phase would consist of simply asking the child to move the robot along a predefined path, as in Task 2. Automatic analysis of the child's gestures during this initial test would allow classifying the user within the taxonomy. With this information on the user's intuitive preferences, the system could give more weight to a specific set of gestures among the ones in the proposed language. For example, if we detect that the user prefers a user-centric viewpoint with full body interactions in the analysis phase, the viewpoint would be set to user-centric while gestures a) and d) would have more relevance than gestures b), e) and f) (see **Fig. 11**). It remains to be studied whether the increase in robot-centric viewpoint preferences continues to rise in children over 12 years old.

In addition, children were not required to use the same gestures in all tasks, and instead they were given freedom to perform or change any gestures they wanted to at all times. As reported in section 3.4.2, some children were observed to change their elicited gestures from one task to another. This was not encouraged nor explicitly mentioned, but rather a spontaneous situation that emerged when they finished Task 1 and started controlling the robot for Tasks 2 or 3. One possible reason for this situation could be that children perceived Task 2 and 3 as a whole, as they had to move the robot over a path. Instead of considering the path as a set of discrete gestures they could give to the robot, they might have envisioned it as a continuous element. This could explain why, instead of using a gesture to indicate each separate action to the robot to follow the path, several children changed their gestures towards a deictic or mimicking approach. In this case, they started pointing the path to the robot as a continuous action, or walked over the path/near the robot to define the directions it should follow.

A very interesting aspect to evaluate would be the relation of the cognitive load for each gesture style. Given the nature of the different gesture styles in terms of its execution process, we can observe these are completely different and hence very difficult to compare within the context of the current experimental design. For example, in some cases, the gesture style entails a continuous movement, while other times the gesture style comprises a very punctual and short movement. This prevents from comparing the different elicited individual gestures based on execution time, as the duration of the gesture would not be a suitable indicator of its difficulty for the participant. Besides, the gesture style, as can be seen in the different tasks of this study, changes due to the nature of the action or goal the participant wants to achieve. For these reasons, this study has rather focused on an evaluation that compares different kind of tasks, comparing the task as a whole, by using time as a comparative measure for age and gender. Nevertheless, it would be very interesting to design a more specific experiment in which different gesture styles could be compared between them in terms of the participants' cognitive load using time as a comparative factor, as well as the participants' perceptions using different kind of gesture styles for the same task.

4.2. Children's perceptions regarding the gestural language

Answering the research question that considered the impact of age and gender in the perceived enjoyment of the robot (RQ5), the results show that there are no differences and that overall enjoyment was rated very highly.

Regarding differences in the perceived easiness of the elicitation task (RQ6), the results showed no differentiation by gender. In terms of age differences, although age group G1 rated this aspect slightly more positively than G2 and G3, the overall assessment was extremely positive. These positive results confirm that the Wizard-of-Oz approach provided the desired effect in the participants, allowing the children to perform the gestures that came most naturally to them. As the children were defining their own gestures, it was also expected that they found it easy to control the robot if the Wizard-of-Oz was correctly carried out. These results will be compared with future results of children using the gesture language, once it has been implemented in the robot. Using these results as a ground base will help to decide whether the implementation of the proposed gesture language suits children's expectations.

The motivation to play was also reported as very positive, without significant differences by gender or age (RQ7). Almost all the children reported they would like to play with the robot again, either at school or outside school. The fact that they would like to play with the robot again even outside school could indicate that they perceived this activity as a playful experience. Powerful interactive scenarios for learning could therefore be developed by making use of this intrinsically motivating technology.

Regarding RQ8 (whether robots can be considered as a promising tool for remote playful activities), the answer is affirmative. The postquestionnaire answers to Q5 shown in **Fig. 16** and **Fig. 17** demonstrate that children are motivated to use the robot in a remotely controlled scenario. This triggered their curiosity towards exploring areas out of their sight, and imagining fun scenarios in which they could use the robot to spy on their classmates, etc. This aspect, in combination with the social component of collaborative playful scenarios with robots, could help to create highly motivating learning activities.

In response to RQ9 (whether robots can be considered as a promising tool for collaborative playful activities), the results of the postquestionnaires showed very positive perceptions (Q6, see **Fig. 18**). This motivates the design and development of collaborative scenarios between children, or even between a child and his parents/teachers. Such scenarios could use the gesture-based control of a robot as a tool towards pursuing a collaborative goal that promotes educational and social values.

4.3. Considerations for gestural language implementation

An interesting aspect that emerged during the study was that the children were so immersed in the activity that they frequently forgot about the tracking technology. In other cases, even though they had been introduced to the tracking device at the beginning of the experiment, several children thought the tracking system was in the robot's camera. Consequently, they sometimes placed themselves not facing the Kinect sensor in order to interact directly with the robot. There is thus a need for interactive spaces with several sensors capable of detecting and responding to variable and natural movements of the children who are not aware of where the tracking mechanisms are while they play.

Comparing the children's elicitation study for drone interaction in this manuscript to previous work in gesture elicitation for drone interaction with adults (Cauchard et al. 2015; Obaid et al. 2016), it can be observed that both children in this study and adults in previous works (Cauchard et al. 2015) tend to consider the drone as an animate being, sometimes even speaking to it as if it was a pet. In addition, children performed a more varied set of gestures to interact with drones than adults in previous studies (Cauchard et al. 2015; Obaid et al. 2016). While adults usually performed gestures using just one or two hands/arms (Cauchard et al. 2015; Obaid et al. 2016), children had a divided preference between those gestures and full body interactions. Adults do not usually move their whole body to mimic the robot's movements, while when designing for children this should be an essential gestural communicative modality.

Given the variability of the gestural language defined by the children, the detection of the elicited gestures stands as a non-trivial computer vision challenge that has to be properly addressed. The Kinect sensor used in our experiments has been discontinued from the market, hence these problems might be studied in a future deployment using different depth-based sensors or the new promised Kinect for Azure v4 sensor. The detection of fine motion interactions within the gestures should be addressed in a way that the system is able to differentiate between different gestures that a framework based on body joints such as Kinect might not be able to distinguish. Hence, the study of computer vision based approaches that combine information from both the color and depth streams for this task seems appropriate. Such information

could then be coupled with learning algorithms capable of capturing distinctive features in the gestures in order to learn how to classify them, such as by using the golden-section search for gesture recognition (Puranam 2005; Wobbrock et al. 2007; Wang and Zhang 2015). Based on the observational analysis of the elicited gestures, these seem to have distinctive features that the proposed heuristics should be able to use in order to differentiate between gestures. Nevertheless, this should be properly assessed in a follow up study that implements the proposed gestural language and the aforementioned heuristics, in order to calculate the precise classification accuracy of the system.

The use of several depth sensors in the same area should also be carefully studied so that each sensor's detection does not interfere with the rest of the devices. These scenarios have already been tackled by previous works (Jones et al. 2013, 2014; Fender et al. 2017), showing feasibility towards the deployment of several sensors in the same room without affecting the detection process.

5. Threats to Validity

Although this study obtained very interesting results, several precautions must be taken before generalizing these results to other contexts.

Despite the fact that the elicitation study is intended to allow children to perform any kind of gesture that feels more natural to them, previous background may condition their responses. For example, children who already own a drone/robot or who have previous experience with videogames, robotics, etc. may have developed their own spatial representations and semiotics. However, this is hard to isolate in current elicitation studies, as children of this age range are now highly exposed to technology. In addition, the experiment involved only one interactive robot. In addition, the experimental Wizard-of-Oz setting, although common in gesture elicitation studies, poses several limitations. Such limitations have to deal with the role of the experimenter, which is very relevant for both in interpreting the gestures of the subjects. In this regard, it is important to assess the children's perception towards the control of the robot, as has been reported in Section 3.4.5, showing no major issues in this regard. The physical attributes of the robot could also have played a role in the specific gestures the children performed. For example, some of the children commented that the robot resembled a pet, and the robot was fitted with a camera that clearly indicated its orientation. Hence, the performed gestures could also have been influenced by these attributes, so that using a different robot, e.g. a spherical one with no marked orientation, may produce a different set of gestures. We believe all these threats to validity could be addressed by performing a new set of experiments to validate the proposed gestural language with different types of ground robots.

6. Conclusions and Future Work

This paper describes the first gesture elicitation study with children using full body interactions to control tangible ground robots. The children's preferred gestures and body language communication with robots in interactive spaces were identified. This has resulted in the definition of a gestural language that accounts for the outlined variabilities of such gestures by age group and gender, which also showed a good acceptance rate in the studied age range. The children were found to be highly motivated towards the activity and the easiness of the interaction was rated as very positive. The next step will be the validation of the proposed gestural language in terms of accuracy of the gesture tracking, usability and learning curves for children with a Kinect sensor. First, the elicited gestural language should be validated by proposing the defined gestures to a different population in order to assess how well this language is accepted by different participant groups. Secondly, the effect of the interactive robot in the elicitation process should also be evaluated by performing similar studies with different robotic devices of diverse characteristics (e.g. dimensions) and variations of their specifications (e.g. different speed, acceleration, interactive responses).

The encouraging results obtained from the questionnaires revealed that interactive spaces with robots using body gestures could become promising collaborative scenarios. We have now started a study in which pairs of children collaborate in controlling a ground mini-robot to achieve a common goal. The results of this study are expected to provide further information on collaborative learning activities with robots using full-body interaction. Moreover, the children's perceptions of the easiness and social aspects of the activity will also be evaluated.

A further aspect that emerged from the postquestionnaires was the children's curiosity about remote interaction. In this respect, we are designing an interactive space for hospitalized children who might not be able to leave their room, such as those in pediatric transplant or oncology wards. Communication and

socialization within the hospital environment are very important aspects to improve children's mental wellbeing during their stay (Lambert et al. 2014), especially in the case of long-term hospitalizations. These long stays at the hospital can produce emotional issues in children, sometimes due to feelings of loneliness or isolation as a result of being separated from their friends or family (Tjaden et al. 2012; Garcia-Sanjuan et al. 2016). Remotely controlling a robot could help these children to explore spaces outside their room or even socialize with other children or adults, thus helping to generate positive emotions during their stay and foster relationships and communication with their teachers/parents/companions. In this regard, animal therapy is often used in hospitals to improve the patients' wellbeing, however children in oncology wards cannot usually benefit from such an experience due to their condition. Several works have studied the combination of animal therapy and games in hospitalized patients, observing their positive effects on morale and stress reduction in kids (Kaminski et al. 2002), and remote interactive applications are being developed to allow children to remotely interact with animals by using tangible robots (Pons et al. 2018). For this reason, this will be an interdisciplinary study in which children could use the proposed gestural language to remotely control a robot to play with an animal located outside the hospital. We believe the intuitive nature of controlling the robot by gestures and the ability to play with an animal in real-time will greatly benefit the mental wellbeing of young patients in these difficult situations.

7.   Acknowledgments

8.   References

Alborzi H, Hammer J, Kruskal A, et al (2000) Designing StoryRooms: Interactive Storytelling Spaces for Children. In: Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '00. ACM Press, New York, New York, USA, pp 95–104

Antle AN, Corness G, Droumeva M (2009) What the body knows: Exploring the benefits of embodied metaphors in hybrid physical digital environments. Interact Comput 21:66–75. doi: 10.1016/j.intcom.2008.10.005

Belpaeme T, Baxter PE, Read R, et al (2013) Multimodal Child-Robot Interaction: Building Social Bonds. J Human-Robot Interact 1:33–53. doi: 10.5898/JHRI.1.2.Belpaeme

Benko H, Wilson AD, Zannier F, Benko H (2014) Dyadic projected spatial augmented reality. In: Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14. pp 645–655

Bobick AF, Intille SS, Davis JW, et al (1999) The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment. Presence Teleoperators Virtual Environ 8:367–391. doi: 10.1162/105474699566297

Bonarini A, Clasadonte F, Garzotto F, Gelsomini M (2015) Blending robots and full-body interaction with large screens for children with intellectual disability. In: Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15. ACM Press, New York, New York, USA, pp 351–354

Cauchard JR, E JL, Zhai KY, Landay JA (2015) Drone & me: An Exploration Into Natural Human-Drone Interaction. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15. ACM Press, New York, New York, USA, pp 361–365

Connell S, Kuo P-Y, Liu L, Piper AM (2013) A Wizard-of-Oz elicitation study examining child-defined gestures with a whole-body interface. In: Proceedings of the 12th International Conference on Interaction Design and Children - IDC '13. ACM Press, New York, New York, USA, pp 277–280

Derboven J, Van Mechelen M, Slegers K (2015) Multimodal Analysis in Participatory Design with Children. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. ACM Press, New York, New York, USA, pp 2825–2828

Dong H, Danesh A, Figueroa N, Saddik A El (2015) An Elicitation Study on Gesture Preferences and

Memorability Toward a Practical Hand-Gesture Vocabulary for Smart Televisions. IEEE Access 3:543–555. doi: 10.1109/ACCESS.2015.2432679

Druin A (1999) Cooperative inquiry: Developing New Technologies for Children with Children. Proc SIGCHI Conf Hum factors Comput Syst CHI is limit - CHI '99 14:592–599. doi: 10.1145/302979.303166

Druin A (2002) The role of children in the design of new technology. Behav Inf Technol 21:1–25. doi: 10.1080/01449290110108659

Druin A, Bederson B, Boltman A, et al (1999) Children as our technology design partners. In: Druin A (ed) The design of children's technology. Morgan Kaufman, San Francisco, CA, pp 51–72

Epps J, Lichman S, Wu M (2006) A study of hand shape use in tabletop gesture interaction. In: CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06. ACM Press, New York, New York, USA, pp 748–753

Fender AR, Benko H, Wilson A (2017) MeetAlive : Room-Scale Omni-Directional Display System for Multi-User Content and Control Sharing. In: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces. pp 106–115

Fernandez RAS, Sanchez-Lopez JL, Sampedro C, et al (2016) Natural user interfaces for human-drone multi-modal interaction. In: 2016 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, pp 1013–1022

Garcia-Sanjuan F, Jaen J, Jurdi S (2016) Towards Encouraging Communication in Hospitalized Children through Multi-Tablet Activities. In: Proceedings of the XVII International Conference on Human Computer Interaction. p 29.1-29.4

Garcia-Sanjuan F, Jaen J, Nacher V, Catala A (2015) Design and evaluation of a tangible-mediated robot for kindergarten instruction. In: Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology - ACE '15. ACM Press, New York, New York, USA, pp 1–11

Gindling J, Ioannidou A, Loh J, et al LEGOsheets: a rule-based programming, simulation and manipulation environment for the LEGO Programmable Brick. In: Proceedings of Symposium on Visual Languages. IEEE Comput. Soc. Press, pp 172–179

Gonzalez B, Borland J, Geraghty K (2009) Whole body interaction for child-centered multimodal language learning. In: Proceedings of the 2nd Workshop on Child, Computer and Interaction - WOCCI '09. ACM Press, New York, New York, USA, pp 1–5

Grønbæk K, Iversen OS, Kortbek KJ, et al (2007) Interactive Floor Support for Kinesthetic Interaction in Children Learning Environments. In: Human-Computer Interaction – INTERACT 2007. Lecture Notes in Computer Science. pp 361–375

Guha ML, Druin A, Chipman G, et al (2005) Working with young children as technology design partners. Commun ACM 48:39–42. doi: 10.1145/1039539.1039567

Hansen JP, Alapetite A, MacKenzie IS, Møllenbach E (2014) The use of gaze to control drones. In: Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14. ACM Press, New York, New York, USA, pp 27–34

Henkemans OAB, Bierman BPB, Janssen J, et al (2017) Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1. Int J Hum Comput Stud 106:63–76. doi: 10.1016/j.ijhcs.2017.06.001

Horn MS, Crouser RJ, Bers MU (2011) Tangible interaction and learning: the case for a hybrid approach. Pers Ubiquitous Comput 16:379–389. doi: 10.1007/s00779-011-0404-2

Hourcade JP (2015) Child computer interaction. CreateSpace Independent Publishing Platform

Höysniemi J, Hämäläinen P, Turkki L (2004) Wizard of Oz prototyping of computer vision based action games for children. In: Proceeding of the 2004 conference on Interaction design and children building a community - IDC '04. ACM Press, New York, New York, USA, pp 27–34

Höysniemi J, Hämäläinen P, Turkki L, Rouvi T (2005) Children's intuitive gestures in vision-based action games. Commun ACM 48:44–50. doi: 10.1145/1039539.1039568

Hsiao H-S, Chen J-C (2016) Using a gesture interactive game-based learning approach to improve preschool children's learning performance and motor skills. Comput Educ 95:151–162. doi: 10.1016/j.compedu.2016.01.005

Jokela T, Rezaei PP, Väänänen K (2016) Using Elicitation Studies to Generate Collocated Interaction Methods. Proc 18th Int Conf Human-Computer Interact with Mob Devices Serv Adjun 1129–1133. doi: 10.1145/2957265.2962654

Jones B, Benko H, Ofek E, Wilson AD (2013) IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13. pp 869–878

Jones B, Shapira L, Sodhi R, et al (2014) RoomAlive: magical experiences enabled by scalable, adaptive projector-camera units. In: Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14. pp 637–644

Kaminski M, Pellino T, Wish J (2002) Play and pets: The physical and emotional impact of child-life and pet therapy on hospitalized children. Child Heal Care 31:321–335. doi: 10.1207/S15326888CHC3104_5

Karam M, Schraefel m. c. (2005) A Taxonomy of Gestures in Human Computer Interactions. Tech. Report, Eletronics Comput. Sci. 1–45

Kistler F, André E (2013) User-defined body gestures for an interactive storytelling scenario. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 8118 LNCS:264–281. doi: 10.1007/978-3-642-40480-1_17

Konda KR, Königs A, Schulz H, Schulz D (2012) Real time interaction with mobile robots using hand gestures. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12. ACM Press, New York, New York, USA, pp 177–178

Kray C, Nesbitt D, Dawson J, Rohs M (2010) User-defined gestures for connecting mobile phones, public displays, and tabletops. In: Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10. ACM Press, New York, New York, USA, pp 239–248

Kurdyukova E, Redlin M, André E (2012) Studying user-defined iPad gestures for interaction in multi-display environment. In: Proceedings of the 2012 ACM international conference on Intelligent User Interfaces - IUI '12. ACM Press, New York, New York, USA, pp 93–96

Lambert V, Coad J, Hicks P, Glacken M (2014) Social spaces for young children in hospital. Child Care Health Dev 40:195–204. doi: 10.1111/cch.12016

Lee S-S, Chae J, Kim H, et al (2013) Towards more natural digital content manipulation via user freehand gestural interaction in a living room. In: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing - UbiComp '13. ACM Press, New York, New York, USA, p 617

Malinverni L, Mora-Guiard J, Pares N (2016) Towards methods for evaluating and communicating participatory design: A multimodal approach. Int J Hum Comput Stud 94:53–63. doi: 10.1016/j.ijhcs.2016.03.004

Mann HB, Whitney DR (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat 18:50–60. doi: 10.1214/aoms/1177730491

Marco J, Cerezo E, Baldassarri S, et al (2009) Bringing Tabletop Technologies to Kindergarten Children. In: Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology. pp 103–111

Michaud F, Caron S (2002) Roball, the Rolling Robot. Auton Robots 12:211–222. doi: 10.1023/A:1014005728519

Micire M, Desai M, Courtemanche A, et al (2009) Analysis of natural gestures for controlling robot teams

on multi-touch tabletop surfaces. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '09. ACM Press, New York, New York, USA, pp 41–48

Mora-Guiard J, Crowell C, Pares N, Heaton P (2016) Lands of Fog: Helping Children with Autism in Social Interaction through a Full-Body Interactive Experience. In: Proceedings of the The 15th International Conference on Interaction Design and Children - IDC '16. ACM Press, New York, New York, USA, pp 262–274

Morris MR (2012) Web on the Wall: Insights from a Multimodal Interaction Elicitation Study. In: Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces. ACM Press, New York, New York, USA, pp 95–104

Morris MR, Wobbrock JO, Wilson AD (2010) Understanding users' preferences for surface gestures. In: Proceedings of Graphics Interface 2010. pp 261–268

Nacher V, Garcia-Sanjuan F, Jaen J (2016) Evaluating the Usability of a Tangible-Mediated Robot for Kindergarten Children Instruction. In: 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT). IEEE, pp 130–132

Nahapetyan VE, Khachumov VM (2015) Gesture recognition in the problem of contactless control of an unmanned aerial vehicle. Optoelectron Instrum Data Process 51:192–197. doi: 10.3103/S8756699015020132

Obaid M, Häring M, Kistler F, et al (2012) User-Defined Body Gestures for Navigational Control of a Humanoid Robot. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp 367–377

Obaid M, Kistler F, Häring M, et al (2014) A Framework for User-Defined Body Gestures to Control a Humanoid Robot. Int J Soc Robot 6:383–396. doi: 10.1007/s12369-014-0233-3

Obaid M, Kistler F, Kasparavičiūtė G, et al (2016) How would you gesture navigate a drone?: a user-centered approach to control a drone. In: Proceedings of the 20th International Academic Mindtrek Conference on - AcademicMindtrek '16. ACM Press, New York, New York, USA, pp 113–121

Pares N, Soler M, Sanjurjo À, et al (2005) Promotion of creative activity in children with severe autism through visuals in an interactive multisensory environment. In: Proceeding of the 2005 conference on Interaction design and children - IDC '05. ACM Press, New York, New York, USA, pp 110–116

Pfeil K, Koh SL, LaViola J (2013) Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles. Proc 2013 Int Conf Intell user interfaces - IUI '13 257--266. doi: 10.1145/2449396.2449429

Piaget J (1973) The child and reality: Problems of genetic psychology. Grossman, New York

Piaget J (1956) The Child's Conception of Space. Norton, New York

Piumsomboon T, Clark A, Billinghurst M, Cockburn A (2013) User-defined gestures for augmented reality. In: CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13. ACM Press, New York, New York, USA, pp 955–960

Pons P, Carrión A, Jaen J (2018) Remote interspecies interactions: Improving humans and animals' wellbeing through mobile playful spaces. Pervasive Mob Comput. doi: 10.1016/j.pmcj.2018.12.003

Puranam MB (2005) Towards Full-Body Gesture Analysis and Recognition. University of Kentucky

Pyryeskin D, Hancock M, Hoey J (2012) Comparing elicited gestures to designer-created gestures for selection above a multitouch surface. In: Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces - ITS '12. ACM Press, New York, New York, USA, pp 1–10

Raffle HS, Parkes AJ, Ishii H (2004) Topobo: a constructive assembly system with kinetic memory. System 6:647–654. doi: 10.1145/985692.985774

Read JC, Macfarlane S, Casey C (2002) Endurability, Engagement and Expectations: Measuring Children's Fun. In: Interaction Design and Children. pp 189–198

Read JC, Markopoulos P (2013) Child–computer interaction. Int J Child-Computer Interact 1:2–6. doi:

10.1016/j.ijcci.2012.09.001

Read JC, Markopoulos P, Parés N, et al (2008) Child computer interaction. In: Proceeding of the twenty-sixth annual CHI conference extended abstracts on Human factors in computing systems - CHI '08. ACM Press, New York, New York, USA, pp 2419–2422

Robins B, Dautenhahn K (2014) Tactile Interactions with a Humanoid Robot: Novel Play Scenario Implementations with Children with Autism. Int J Soc Robot 6:397–415. doi: 10.1007/s12369-014-0228-0

Robins B, Dautenhahn K, Te Boekhorst R, Nehaniv CL (2008) Behaviour delay and robot expressiveness in child-robot interactions: a user study on interaction kinesics. Proc 3rd ACMIEEE Int Conf Hum Robot Interact 17–24. doi: 10.1145/1349822.1349826

Ruiz J, Li Y, Lank E (2011) User-defined motion gestures for mobile interaction. In: Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11. ACM Press, New York, New York, USA, p 197

Rust K, Malu M, Anthony L, Findlater L (2014) Understanding childdefined gestures and children's mental models for touchscreen tabletop interaction. In: Proceedings of the 2014 conference on Interaction design and children - IDC '14. ACM Press, New York, New York, USA, pp 201–204

Salter T, Dautenhahn K, Te Boekhorst R (2006) Learning about natural human-robot interaction styles. Rob Auton Syst 54:127–134. doi: 10.1016/j.robot.2005.09.022

Sanghvi J, Castellano G, Leite I, et al (2011) Automatic analysis of affective postures and body motion to detect engagement with a game companion. In: Proceedings of the 6th international conference on Human-robot interaction - HRI '11. ACM Press, New York, New York, USA, pp 305–311

Sanna A, Lamberti F, Paravati G, Manuri F (2013) A Kinect-based natural interface for quadrotor control. Entertain Comput 4:179–186. doi: 10.1016/j.entcom.2013.01.001

Sato E, Yamaguchi T, Harashima F (2007) Natural interface using pointing behavior for human-robot gestural interaction. IEEE Trans Ind Electron 54:1105–1112. doi: 10.1109/TIE.2007.892728

Schaper M-M, Malinverni L, Pares N (2015) Sketching through the body: Child-generated gestures in Full-Body Interaction Design. In: Proceedings of the 14th International Conference on Interaction Design and Children - IDC '15. ACM Press, New York, New York, USA, pp 255–258

Schaper M-M, Pares N (2016) Making Sense of Body and Space through Full-Body Interaction Design. In: Proceedings of the The 15th International Conference on Interaction Design and Children - IDC '16. ACM Press, New York, New York, USA, pp 613–618

Seyed T, Burns C, Costa Sousa M, et al (2012) Eliciting usable gestures for multi-display environments. In: Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces - ITS '12. ACM Press, New York, New York, USA, p 41

Shimon SSA, Morrison-Smith S, John N, et al (2015) Exploring User-Defined Back-Of-Device Gestures for Mobile Devices. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '15. ACM Press, New York, New York, USA, pp 227–232

Sipitakiat A, Nusen N (2012) Robo-blocks: a tangible programming system with debugging for children. In: Proceedings of the 11th International Conference on Interaction Design and Children - IDC '12. ACM Press, New York, New York, USA, p 98

Soler-Adillon J, Ferrer J, Pares N (2009) A novel approach to interactive playgrounds: The interactive slide project. In: Proceedings of the 8th International Conference on Interaction Design and Children - IDC '09. ACM Press, New York, New York, USA, pp 131–139

Stiefelhagen R, Fogen C, Gieselmann P, et al (2004) Natural human-robot interaction using speech, head pose and gestures. In: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566). IEEE, pp 2422–2427

Subrahmanyam K, Greenfield PM (1994) Effect of video game practice on spatial skills in girls and boys. J Appl Dev Psychol 15:13–32. doi: 10.1016/0193-3973(94)90004-3

Sugiyama J, Tsetserukou D, Miura J (2011) NAVIgoid: Robot navigation with haptic vision. SIGGRAPH Asia 2011 Emerg Technol SA'11 15:4503. doi: 10.1145/2073370.2073378

Takahashi T, Morita M, Tanaka F (2012) Evaluation of a tricycle-style teleoperational interface for children: A comparative experiment with a video game controller. In: 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. IEEE, pp 334–338

Tanaka F, Takahashi T (2012) A tricycle-style teleoperational interface that remotely controls a robot for classroom children. In: Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction - HRI '12. ACM Press, New York, New York, USA, pp 255–256

Tjaden L, Tong A, Henning P, et al (2012) Children's experiences of dialysis: A systematic review of qualitative studies. Arch Dis Child 97:395–402. doi: 10.1136/archdischild-2011-300639

Vatavu R-D (2017) Smart-Pockets: Body-deictic gestures for fast access to personal data during ambient interactions. Int J Hum Comput Stud 103:1–21. doi: 10.1016/j.ijhcs.2017.01.005

Vatavu R-D (2012) User-defined gestures for free-hand TV control. In: Proceedings of the 10th European conference on Interactive tv and video - EuroiTV '12. ACM Press, New York, New York, USA, pp 45–48

Vatavu R-D, Wobbrock JO (2015) Formalizing Agreement Analysis for Elicitation Studies: New Measures, Significance Test, and Toolkit. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15. ACM Press, New York, New York, USA, pp 1325–1334

Vatavu R-D, Wobbrock JO (2016) Between-Subjects Elicitation Studies: Formalization and Tool Support. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16. ACM Press, New York, New York, USA, pp 3390–3402

Voyer D, Voyer S, Bryden MP (1995) Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. Psychol Bull 117:250–270. doi: 10.1037/0033-2909.117.2.250

Wainer J, Robins B, Amirabdollahian F, Dautenhahn K (2014) Using the Humanoid Robot KASPAR to Autonomously Play Triadic Games and Facilitate Collaborative Play Among Children With Autism. IEEE Trans Auton Ment Dev 6:183–199. doi: 10.1109/TAMD.2014.2303116

Wang Y, Zhang L (2015) A Track-based Gesture Recognition Algorithm for Kinect. Appl Mech Mater 738–7399:334–338. doi: 10.4028/www.scientific.net/AMM.738-739.334

Wilson AD, Benko H (2010) Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In: Proceedings of the 23nd annual ACM symposium on User interface software and technology - UIST '10. ACM Press, New York, New York, USA, pp 273–282

Wobbrock JO, Hall MG, Wilson AD (2007) Gestures without Libraries , Toolkits or Training : A $ 1 Recognizer for User Interface Prototypes. In: Proceedings of the 20th annual ACM symposium on User interface software and technology. pp 159–168

Wobbrock JO, Morris MR, Wilson AD (2009) User-defined gestures for surface computing. In: Proceedings of the 27th international conference on Human factors in computing systems - CHI 09. ACM Press, New York, New York, USA, pp 1083–1092