

# Evaluating the Child-Robot Interaction of the NAOTherapist Platform in Pediatric Rehabilitation

José Carlos Pulido · José Carlos González · Cristina Suárez-Mejías · Antonio Bandera · Pablo Bustos · Fernando Fernández

Received: date / Accepted: date

**Abstract** NAOTherapist is a cognitive robotic architecture whose main goal is to develop non-contact upper-limb rehabilitation sessions autonomously with a social robot for patients with physical impairments. In order to achieve a fluent interaction and an active engagement with the patients, the system should be able to adapt by itself in accordance with the perceived environment. In this paper, we describe the interaction mechanisms that are necessary to supervise and help the patient to carry out the prescribed exercises correctly. We also provide an evaluation focused on the child-robot interaction of the robotic platform with a large number of schoolchildren and the experience of a first contact with three pediatric rehabilitation patients. The results presented are obtained through questionnaires, video analysis and system logs, and have proven to be consistent with the hypotheses proposed in this work.

---

J. C. Pulido (first author) and J. C. González (second author) contributed equally to this work.

---

J. C. Pulido (✉) · J. C. González · F. Fernández  
Computer Science and Engineering  
Universidad Carlos III de Madrid  
Av. de la Universidad 30, 28911, Madrid, Spain  
Tel.: +34-91-6245981  
E-mail: {jcpulido (✉), josgonza, ffernand}@inf.uc3m.es

C. Suárez-Mejías  
Hospital Universitario Virgen del Rocío  
Av. Manuel Siurot, s/n, 41013 Sevilla, Spain  
E-mail: cristina.suarez.exts@juntadeandalucia.es

A. Bandera  
Universidad de Málaga  
Campus de Teatinos, s/n, 29071 Málaga, Spain  
E-mail: ajbandera@uma.es

P. Bustos  
Robolab, Universidad de Extremadura  
Campus Universitario, s/n, 10071 Cáceres, Spain  
E-mail: pbustos@unex.es

**Keywords** Social Human-Robot Interaction · Rehabilitation Robotics · Socially Assistive Robotics · Control Architectures and Programming · Automated Planning

## 1 Introduction

Socially Assistive Robotics (SAR) is a growing field whose purpose is to use robots to undertake certain social needs. This term represents all those robotic platforms that provide a service or assistance to people through social interaction [13]. In the last ten years, a wide variety of assistive devices have been developed as support systems and many of them have gained far-reaching acceptance among users and professionals alike [30]. This has opened up new lines of research in different application domains, including physical and cognitive rehabilitation.

Traditional methods of physical rehabilitation comprise continuous repetitions of movements according to the clinical conditions of the patient. This can bring about a loss of interest and reduced therapy engagement on the part of the patient (especially children). Consequently, the therapists need more time and effort when carrying out the therapy sessions.

Our proposed system is called NAOTherapist and it is the result of a new development phase in the Therapist project [5]. In the first approach, a bear-like robotic platform called Ursus executed a sequence of preprogrammed behaviors to carry out rehabilitation movements with the upper limbs [38]. This and most other SAR approaches still overlook the autonomy and quick response of the robot which are essential points of SAR platforms. We consider that during rehabilitation sessions, the lack of human intervention and a fluent inter-

action promotes an active engagement with and commitment of the patients, in which the robot captures the full attention by being prominent in the room. We have taken all these elements into account in designing the NAOTherapist architecture and use case [20]. In essence, the use case that we are considering in this work consists of a NAO robot which performs a set of prescribed arm-poses that a patient has to imitate. The system is able to react autonomously and check the pose of the patient helping him to correct it, if required<sup>1</sup>. This automatic reasoning is carried out using Automated Planning techniques [19], where the perceived environment is encoded as a symbolic representation of the state of the world using the standard Planning Domain Definition Language (PDDL) [15]. This is briefly explained in Section 3.

In pediatric rehabilitation, patients are children who need constant motivational reinforcement from the therapists and a great variety of activities. Our robotic platform focuses on upper-limb motor rehabilitation for patients that suffer from cerebral palsy and obstetric brachial plexus palsy. The biggest challenge is to ensure that the patients are committed and follow the prescribed treatment closely. So, proving that the NAOTherapist platform is able to achieve an active engagement with patients in pediatric rehabilitation is required.

In order to understand the philosophy of the interaction that this work pursues, the mechanisms associated with perception, interaction, action and monitoring are described in Section 5. The rest of the document presents the evaluation setup that has been designed from six established hypotheses to be demonstrated (see Section 6). Two different scenarios and users have been selected: on the one hand, a large number of healthy children in schools to determine the degree of engagement in the activity together with the autonomy of the robotic system. On the other hand, three selected pediatric patients from the *Hospital Universitario Virgen del Rocío* (HUVR) of Seville have a first experience with the robotic tool and share their impression of the usefulness of the NAOTherapist prototype. The evaluation mechanisms are based on questionnaires to participants, relatives and experts, interaction level from video analysis and logs of the vision-action system. The results of this paper seek to demonstrate the potential of these novel robotic tools in the area of pediatric rehabilitation, where a social robot is an extra motivational component to facilitate the development of these tedious treatments. Next Section 2 summarizes the main related work.

## 2 Related Work

The development of new devices to support neurological recovery is a current challenge for clinical professionals and engineers [39, 32]. Particularly, in the last decade robotic applications have demonstrated their great potential as novel approaches [9, 3]. These devices represent those robots that provide a service or assistance to people. Following the taxonomy provided by Feil-Seifer and Mataric of social robotics [13], three main categories can be identified:

**Socially Interactive Robotics (SIR)** comprise those robots whose main task is based on social interaction [14]. Their purpose is not necessarily to be of assistance to the user. Robotic butlers and entertainment robots are clear examples [21, 28].

**Assistive Robotics (AR)** provide assistance to people with no social interaction. For instance, *wearable* robots or exoskeletons for patients with spinal cord injuries increase the range of movements, thus improving their motor skills [35]. Advanced mobility aides are also developed for elderly and visually impaired people as well [34, 10, 25]. There are also robotic platforms that aim to rehabilitate an affected limb by carrying out movements with a controlled resistance [4, 23] and others combine virtual games with remote control techniques for the same purpose [37]. Robot-Mediated Therapy (RMT) devices are available for children which “wears” the patient’s body driving their joints during the rehabilitation process [6, 18, 31].

**Socially Assistive Robotics (SAR)** is the intersection of AR and SIR. This category includes robots that provide assistance through social interaction [38, 12, 7, 17], where NAOTherapist is located. Current trends of SAR seek to accomplish their goals with no physical interaction with the patient [11]. These robots should be able to move autonomously in human environments, interact and socialize with people. Testing and deploying a SAR platform reduces the safety risk, since it is based on non-contact human-robot interaction. The success of these approaches is given by the emotional bonds between the patient and the robot, improving the motivation to continue with the treatment [29, 8, 40, 2, 24]. These platforms must deal with a number of challenges [39, 13]. On the one hand, a SAR system must really satisfy the needs for it was intended. In other words, these robots must be able to perceive the environment and react accordingly. Otherwise the system may be ineffective at achieving measurable improvements in rehabilitation therapies. A higher level of autonomy implies less human intervention, saving time and effort. On the other hand, verbal and non-verbal communication, voice, feedback and physical appear-

<sup>1</sup> Video of the NAOTherapist use case: <https://youtu.be/75xb39Q8QEg>

ance are key points in catching the attention of patients and ensuring a fluent interaction.

There are many SAR approaches with different degrees of success and sophistication. A modern approach for stroke patients is the uBot-5 robot which aims to drive upper-limb physical exercises combined with speech therapy [7]. The platform is a humanoid robot, 86 cm tall and 16 kg in weight with speakers and a screen in place of the head where pre-recorded videos and animations of human faces can be reproduced to provide social stimuli. Each arm has 4 degrees of freedom but lacks mobile hands. An expert must teleoperate the robot during sessions. The robot carries out movements to be followed by the patient and gives clues in the speech therapy, but all the results need to be recorded by the experts to evaluate the progression of the patient. Thus, it does not save the time of professionals, who are still necessary to supervise and control the whole therapy.

KindSAR [16] use a NAO robot to promote the development of children through social interaction and explore the relationship between performance and engagement. The interaction is evaluated using video data from only 11 children, which may not be a sufficiently representative population.

### 3 NAOTherapist Architecture

The components of the NAOTherapist architecture have been designed using the RoboComp framework [27], which has a development environment, tools and reusable components to control robotic platforms. Each RoboComp component is connected to the others using the Internet Communications Engine (Ice) framework through TCP/IP. The transmission of the data is independent of the language in which the components have been programmed because they use shared Ice interfaces. In our architecture, we have reused one RoboComp component to control a Microsoft Kinect 3D sensor. It uses the Kinect for Windows SDK to serve the human body characteristics to the rest of the components. The whole NAOTherapist architecture is structured in three levels of planning [20]:

**High-level planning** is a search-and-selection task addressed using Automated Planning by a component called Therapy Designer [36]. All exercises available in the knowledge base are considered, but only a set of them are included in a session, thus preserving the variability. The planning process is carried out by a Hierarchical Task Network (HTN) algorithm [33]. If there are no exercises available to plan a therapy, this model is able to suggest new exercises whose attributes comply with the established requirements and medical criteria.

**Medium-level planning** refers to the execution of the planned sessions individually, reacting in accordance with the environment perceived by a Kinect device and the sensors of the robot. A Decision Support component is controlled by the PELEA architecture [1] which is in charge of planning and monitoring the execution of the exercises and, if required, making decisions with respect to an unexpected perceived state. The knowledge is modeled as a classical planning domain in PDDL (Planning Domain Definition Language) [15] considering the set of actions that the robot can perform in each session and possible unexpected situations. In this way, the robotic platform is able to behave autonomously as described in Section 5.

**Low-level planning** comprises the decomposition of medium-level actions into a set of instructions that are executed by the robot. For instance, moving the arms to a certain pose, changing the eye color, showing animations, etc. At this level the path planner of the robot performs a planning process to move its joints by estimating the trajectories.

It should be pointed out that the goal of this paper is to evaluate the child-robot interaction which mainly relies on medium-level planning. Therefore, the next sections describe the main elements of this level in depth.

### 4 Perceiving the State of the World

The state of the world is an abstraction of the environment in which the robot works. This is modeled as a classic PDDL automated planning problem and describes the environment using predicates and functions. Some of these predicates control transitions between actions and are only changed internally by the effects of the planned action; but others are changed by external events (exogenous predicates). For instance, the values of the predicates *patient\_detected* and *correct\_pose* are obtained externally from the sensors. The recreation of the actual state of the world requires data to be captured from the sensors and to infer visual information in the Vision component to decide the value of the exogenous predicates.

The Vision component provides a set of methods to the Executive component, in order to return the externally-processed information captured by the Kinect Sensor component. These methods address the following two aspects: pose comparison and situation awareness.

**Pose comparison** uses an estimation of the anthropometric model of the user provided by the Kinect Sensor component and calculates the angles between joints with respect to the anatomical planes for each

arm. The system stores each pose in a knowledge base as static 3D skeletons to compare them with the ones provided by the 3D sensor and to move the robot accordingly. Then, the method calculates the difference between the joints of the desired pose and the patient's performed in terms of the normalized Euclidean distance. Given the angles from joints  $a_i$  where  $i = 1...4$  and  $a_i \in \{\text{shoulder rotation, shoulder opening, elbow rotation and elbow opening}\}$ , the distance  $d(a^h, a^r)$ , where  $h$  refers to human and  $r$  to robot, is computed and normalized between 0 to 1 following Equation 1.

$$d(a^h, a^r) = 1 - \left( \frac{1}{1 + \sqrt{\sum_{i=1}^4 (a_i^h - a_i^r)^2}} \right) \quad (1)$$

Given  $d(a^h, a^r)$ , the pose of the human is considered correct if the  $d(a^h, a^r)$  per each arm is less than a dynamic threshold  $\theta$  and is incorrect otherwise.

It is important to note that a pose is accepted if it is maintained for a determined amount of time. The duration of a pose is established by the therapist according to the configuration of the exercise, so several comparisons are needed in order to accept a pose or not. There will be one comparison per received video frame. When the system is checking the pose, it takes and compares as many video frames with 3D skeleton data as the system can handle, as can be seen in Algorithm 1. The greater amount of samples, the more accurate check result. This explains the need of having a fast-to-calculate equation (Equation 1) to determine a correct pose.

Firstly, before starting to measure the duration of the pose, the system waits a maximum of 4 seconds for the patient to pose correctly. This requires 3 consecutive valid comparisons to avoid possible false positives with the 3D sensor. When the patient starts the pose correctly, the system triggers the timer for the pose and carries out as many comparisons as possible, counting failures and successes. Finally, the pose is accepted if the number of failures is less than the 20% throughout the total duration of the pose. In the case that the pose is incorrect, the function `getLastIncorrectJoints()` returns the last three comparisons to determine the limb or limbs to be corrected (left, right or both), giving the appropriate verbal feedback.

The “dynamic-comparison threshold”  $\theta$  takes values from 0.28 to 0.4, which have been determined experimentally by the therapists. The minimum represents the strictest value to be compared with  $d(a^h, a^r)$ , so a more accurate imitation will be needed, while the maximum is the most permissive. In every session,  $\theta$  is initialized to 0.28 and is updated after evaluating the success of the patient throughout each pose. As can be seen in Algorithm 2, the system allows three attempts

---

**Algorithm 1: Check Pose**


---

```

Input: Pose, Duration, Threshold
Data: MaxTimeToStart, MinCompsToStart,
        MaxFailProportion
Output: Checking result
// 1st: Waiting first correct comparisons
1 EndTime ← MaxTimeToStart+CurrentTime();
2 NumCompsOk ← 0;
3 while NumCompsOk < MinCompsToStart and
  IsPatientReady() and CurrentTime() < EndTime do
4   Comparison ← CompareCurrentPose(Pose);
5   RobotSetEyeColor(Comparison, Threshold);
6   if isValid(Comparison, Threshold) then
7     NumCompsOk ← NumCompsOk+1;
8   else
9     NumCompsOk ← 0;
10 if CurrentTime() < EndTime then
11   return PatientNotReady;
12 if NumCompsOk < MinCompsToStart then
13   return GetLastIncorrectJoints();
// 2nd: Checking throughout the pose duration
14 EndTime ← Duration+CurrentTime();
15 NumCompsOk ← 0;
16 NumCompsFail ← 0;
17 while IsPatientReady() and CurrentTime()
  < EndTime do
18   Comparison ← CompareCurrentPose(Pose);
19   RobotSetEyeColor(Comparison, Threshold);
20   if isValid(Comparison, Threshold) then
21     NumCompsOk ← NumCompsOk+1;
22   else
23     NumCompsFail ← NumCompsFail+1;
// 3rd: Returning results
24 if CurrentTime() < EndTime then
25   return PatientNotReady
26 NumCompsTotal ← NumCompsOk+NumCompsFail;
27 if NumCompsFail/NumCompsTotal
  > MaxFailProportion then
28   return GetLastIncorrectJoints();
29 else
30   return PoseOk;

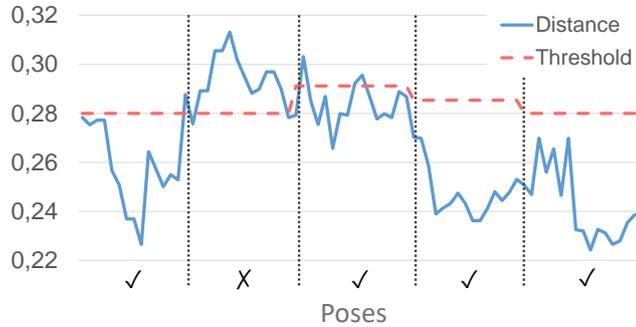
```

---

(with two different correction types) to carry out a pose correctly, otherwise it is omitted. In this case,  $\theta$  is increased by 4%. In contrast, when the patient performs a pose correctly at the first attempt, the threshold is decreased by 2%. These percentages determine the speed of the evolution of  $\theta$ , but always respecting the limits of the threshold.

Figure 1 shows an example of the update of  $\theta$  depending on the values of  $d(a^h, a^r)$  throughout 5 consecutive poses. For clarity, in this example there is only one try per pose. The first pose is correct since less than 20% of the calculated distances are over the threshold. However, it is not decreased because its value is the minimum. The second one is incorrect, so the threshold is increased by 4% for the next pose. The third pose

would have been incorrect if the threshold had not been increased. This and the last two poses are correct so the threshold is decreased by 2% each one until reaching the minimum again.



**Fig. 1** Example of the evolution of the dynamic-comparison threshold according to the calculated distance  $d(a^h, a^r)$  for each processed video frame throughout 5 consecutive poses.

The capabilities of patients can differ widely, so it is necessary to customize the level of difficulty while training for rehabilitation purposes. This explains how the system behaves by being more permissive or not according to the performance and success of the patient during the session. The pose comparison values and threshold are also used to change the color of the eyes of the robot from red to green according to the correctness of the pose.

The limits of  $\theta$  were estimated during evaluation sessions in which therapists labeled several postures as correct or incorrect to determine the average values of the minimum and the maximum. In the same way, the update percentages of  $\theta$  were established experimentally by the therapists to find a suitable speed of the evolution of the threshold for the targeted patients. Although currently the same values are used for every patient, it is planned to have a customized set of constants in a future work.

The comparison made for each received video frame throughout the duration of the pose and the use of the dynamic threshold allow both the patient and 3D sensor to have enough margin of failures and inaccuracies without compromising a fluent interaction. We assume that the majority of the detection errors can be absorbed by this battery of consecutive comparisons.

**Situation awareness** refers to those situations that can appear during sessions and are taken into account in our model. All situations considered can be included in the deliberative model using the Vision component to act accordingly. For instance, if the patient leaves the training area, sits down or stops doing the exercises.

---

### Algorithm 2: Execute Pose

---

**Input:** Pose, Duration  
**Data:** Threshold  
**Output:** Execution result

```

1 Failures  $\leftarrow$  0;
2 Accepted  $\leftarrow$  False;
3 while Failures < 3 and not Accepted do
4   RobotBehavior(Pose);
5   Check  $\leftarrow$  CheckPose(Pose, Duration, Threshold);
6   if Check = PatientNotReady then
7     RobotBehavior(PatientNotReady);
8   else if Check = PoseOk then
9     RobotBehavior(PoseOk);
10    UpdateThreshold(Failures);
11    Accepted  $\leftarrow$  True;
12  else if Failures = 0 then
13    Failures  $\leftarrow$  1;
14    RobotBehavior(NormalCorrect(Pose, Check));
15  else if Failures = 1 then
16    Failures  $\leftarrow$  2;
17    RobotBehavior(MirrorCorrect(Pose, Check));
18  else
19    Failures  $\leftarrow$  3;
20    RobotBehavior(PoseSkipped);
21    UpdateThreshold(Failures);
22 return Accepted;
```

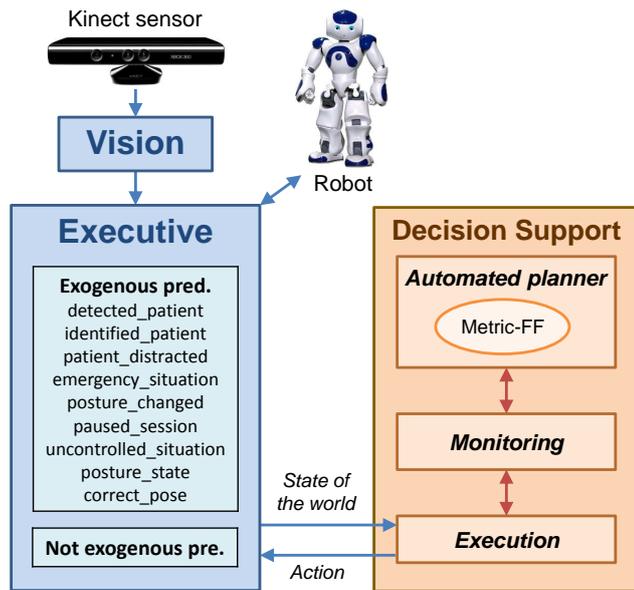
---

## 5 Session Monitoring and Execution

This section explains the reasoned deliberation of medium-level actions according to the perceived environment. Five components of the architecture are involved in this task: Decision Support, Executive, Vision, Kinect Sensor and Robot, as shown in Figure 2.

In essence, the Executive component manages the control of a session and executes the medium-level planned actions. For this purpose, this module communicates with the Decision Support component, the Vision component and the Robot component. The Executive does not take any decision on the next action to be executed by the robot, since this task belongs to Decision Support. When the system has finished the last action, the Executive component asks for the next action from Decision Support. To do so, the Executive needs an accurate enough representation of the environment in which the robot is operating. This is called the “state of the world” (Figure 2). This state of the world is sent to the Decision Support to plan the following actions needed to finish the session.

The Executive component is responsible for maintaining an updated state of the world, requesting the required information from the Vision and Robot, as shown in Figure 2. The Executive has the actual state of the world obtained through the sensors, and Decision Support has the expected state of the world generated internally through the effects of the planned actions.



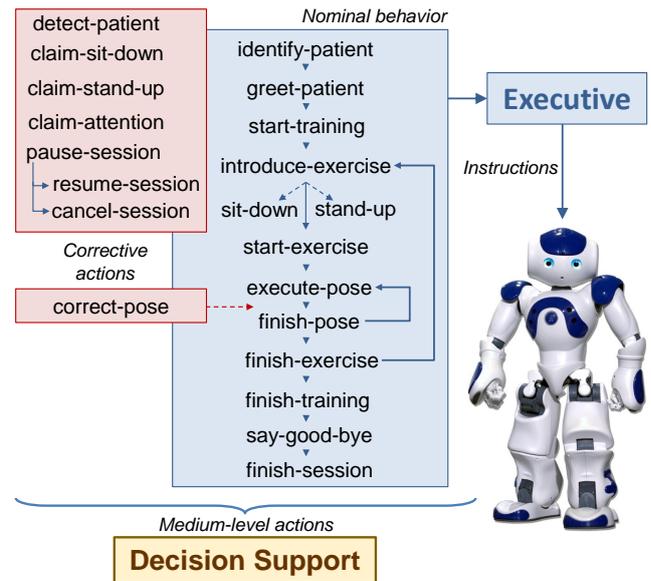
**Fig. 2** Execution flow of medium-level planning with the PELEA sub-architecture embedded into the Decision Support component.

When these states differ in some predicate, the previous plan is invalidated and Decision Support finds a new one from the actual state and then returns the new next action. This is called the replanning process. It is controlled by the PELEA architecture [1] which is integrated into the Decision Support component. When the actual state of the world is the same as the expected one, the next action in the previous plan is returned by the Decision Support without the need to replan. The Monitoring module of PELEA makes a comparison of both states and executes the Metric-FF planner [22] to generate a new plan only when it is needed.

### 5.1 Medium-level Actions

The Executive component controls which behavior is triggered for each action received from the Decision Support (Figure 3). Some actions are simply to control the planning process, but others require the use of sensors, movements of the robot, speech, etc. The planning follows a nominal behaviour, without considering unexpected events. When one of these situations happens, a replanning is triggered and certain corrective actions are planned in order to return to the nominal behavior flow. The list of all possible actions and their interpretation by the Executive component is detailed below:

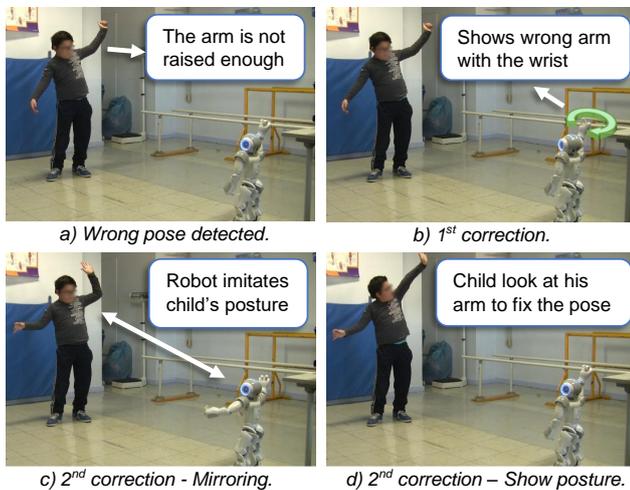
- detect-patient: The execution always starts with this action. It asks the Vision component if there is a person in front of the sensor.



**Fig. 3** Flowchart of the nominal behavior of an initial planning, along with corrective actions that could take place in further replannings. Each possible action is translated into generic instructions to the robot.

- identify-patient: The system loads the respective patient's profile.
- greet-patient: The robot gives the patient a wave and plays a greet message.
- start-training: The robot introduces the ongoing activity to the child.
- introduce-exercise: The robot gives a short explanation of the next exercise before starting it. The corresponding speech is obtained from the knowledge base of exercises.
- stand-up: The robot stands up.
- sit-down: The robot sits down.
- start-exercise: It restarts all pose counters and timers to prepare the system for the upcoming exercise.
- execute-pose: This is one of the most important actions. The Executive component sends to the robot the pose to be imitated with both arms. The robot is in charge of planning the movement interpolation at a low level. Each pose is maintained as long as indicated in the exercise. If the patient is able to hold the pose for the required time, it is considered as correct in the state of the world.
- correct-pose: It is executed if the last pose has not been performed correctly or has not been maintained for the required amount of time. When comparing the pose, the Vision component gives an array of numbers to the Executive which indicates how much the patient has deviated from the expected pose. Based on these numbers, the dynamic-comparison threshold value (explained in Section 4)

and the current attempt, the Executive component starts the correction mechanism (Figure 4). In the first correction, the robot twists the wrist of the incorrect arm or arms and tells the child that the pose must be corrected. In the second correction, the robot imitates the detected posture of the patient, approximately, and shows him how to move the arms to achieve the correct pose. This is called “mirrored correction”. Algorithm 2 describes when to carry out each correction. These two mechanisms provide helpful feedback to users and help them to get closer to the correct pose. If the patient fails these two corrections, the pose is omitted.



**Fig. 4** Pose-correction procedure: first correction (standard) and second correction (mirrored).

- finish-pose: It prepares the system for the upcoming pose.
- finish-exercise: The robot tells the patient that they have finished the current exercise.
- finish-training: The robot wipes out imaginary sweat from its brow while says that he is tired, and informs the patient that the training is finished for today.
- perform-relaxation: The robot takes a break between exercises and encourages the child to breathe deeply for recovery. For this, the robot executes an animation in which it opens its arms, plays inhalation and exhalation sounds and simulate the closing of its eyes by turning off the ring of LEDs of the eyes progressively.
- say-good-bye: The robot waves the patient good-bye.
- finish-session: The robot sits down, starts sleeping and waits for the next patient.
- claim-stand-up: If the patient is seated and the exercise requires him to be standing, the robot asks the patient to stand up.
- claim-sit-down: If the patient is standing and the exercise requires him to be seated, the robot asks the patient to sit down.
- claim-attention: If the Vision component detects that the patient is distracted, the robot attracts his attention.
- pause-session: The session is paused, so the therapist must check why. The system waits until the therapist resumes the execution or cancels the session.
- resume-session: This is triggered by the therapist using the user interface to remove the PDDL predicate that pauses the session and to continue with the rehabilitation.
- cancel-session: This is triggered by the therapist using the user interface to cancel the session. The robot sits down and goes to sleep to wait for another patient.

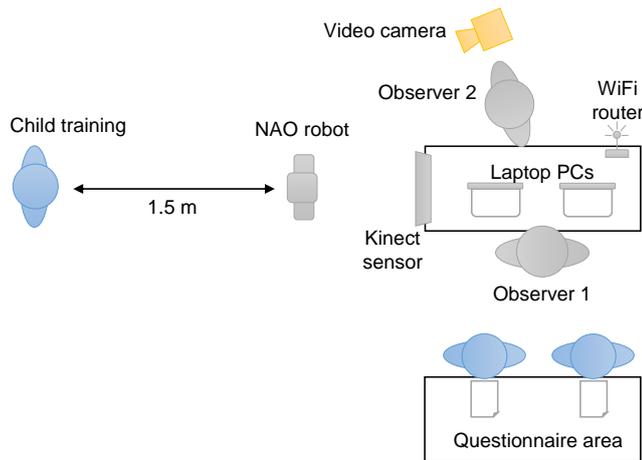
## 6 Experimental Design

We have made two main types of evaluation. The first type was carried out with 117 healthy children from two schools. All participants were volunteers that speak Spanish as their first language with ages between 5 and 9 years old (more details later in Table 2). NAOTherapist was presented as an educational activity about robotics in the school. The main objective of this evaluation was to analyze the child-robot interaction and solve incoming technical issues. The architecture was improved after each experiment to prepare a polished version for the second type of evaluation that was made in the HUVR with 3 patients with upper-limb motor impairments. The main objectives were to evaluate the performance of the overall architecture in a real-case scenario and the children’s reactions using NAOTherapist as a rehabilitation support tool.

These are not long-term experiments, but they allow our objectives to be evaluated at this development stage: the autonomy of the robotic platform, the quality of the child-robot interaction, and the ability of the robotic framework to engage the children throughout the therapy. All data was extracted using application logs, questionnaires, video annotations and the observers’ comments.

## 6.1 Procedure Design

All evaluations in schools share the same setup (Figure 5). Before interacting with the robot, the participants had a first contact with NAO. They can see its appearance, features and some basic skills, but the child does not know exactly how the therapy session works. Then, the child is accompanied to the experimental room and he waits in front of the robot, until the activity starts.



**Fig. 5** Experimental setup for the schoolchildren evaluations.

The use case starts when the child enters in the experimental room and finds the robot seated and “sleeping” at around 1.5 meters from him. Then, the system carries out the appropriate actions one by one to establish the session. These actions have been explained in Section 5.1. NAO starts blinking and wakes up greeting the child and explains how they are going to do exercises together with the arms. Then, they train using the different exercises in the evaluation: 2 for schoolchildren and 4 for pediatric patients. When the training finishes, the robot wipes sweat from his brow, congratulates the child, says good-bye and goes to sleep again. Finally, the children fill a questionnaire whose results are detailed later in Section 7.1. The session is closely observed by two researchers without interfering in the process since it works autonomously until the end. The children could ask any question to the observers in order to answer the questions as correctly as possible.

Robotic rehabilitation therapy sessions involve several problems which are addressed by the NAOTherapist architecture such as RGBD human pose detection, inverse kinematics and task planning and replanning. In the evaluation, the exercises come from real activities used in the hospital to rehabilitate children with

these disabilities. The poses showed by the robot have been designed by the clinical experts taking into account these two criteria: the poses should be detectable by the 3D Kinect sensor and should be also executable by the NAO robot. This means that our system has two limitations that every professional must consider, the first is because of the detectable poses of the Kinect 3D sensor and the second because of the pose compatibility with the joints of the NAO robot.

## 6.2 Hypotheses

The experiments of these evaluations aim to validate the following hypotheses:

- **H1.** “Children are engaged with the therapy and make an effort to follow the session with the robot”.
- **H2.** “Children like to do the exercises with the robot”.
- **H3.** “Children consider the robot as a social and friendly entity”.
- **H4.** “Children are able to carry out the rehabilitation session without previous explanations”.
- **H5.** “The robot is able to carry out the session autonomously and fluently”.
- **H6.** “Experts of the hospital consider that the robot is a useful clinical support tool for rehabilitation”.

## 6.3 Measurements and Metrics

In order to validate the proposed hypotheses, we use three evaluation mechanisms: questionnaires, analysis of the video data and application logs.

The questions in the questionnaires have only two or three possible options. This was recommended by the therapists consulted because it is clearer for young children to have few options to reply. Statements of the children’s questionnaire are included in Appendix A. In the following, almost all of the results of the questionnaires are presented with a value of between 0 and 1, being 1 the most desirable option for us. For the evaluation in the hospital we also provide a questionnaire for the observers (family, physicians and therapists) which is detailed in Appendix B.

In the children’s questionnaire, they also have to select five adjectives from a list which they think are better to describe the robot. These adjectives are classified to measure their perception of the robot as a social entity, instead of an artificial one. Social adjectives like friendly or angry increase the score (+2 for good ones or +1 for bad) and other adjectives for artificial entities like artificial or delicate decrease the score (-1 for good ones or -2 for bad). We have a balanced list

of 8 social and 8 artificial adjectives. The social vs. artificial perception metric can take values from -9 to 9. The questionnaire system has been adapted from the Therapist project [5].

The sessions of the last 50 schoolchildren share the same set of exercises, forming a very homogeneous group to analyze their video data. We used annotations with continuous duration values in accordance with Table 1. The quantitative evaluation of these annotations allows the reactions of the child to be classified on four different aspects of interaction: emotions during the session, effort and attitude while performing the activities, the child’s gaze and the communication with the robot. Each aspect has a track of annotations indicating the corresponding behavior at every moment.

**Table 1** Coding Scheme for Video Annotation

Aspect	Score	Behavior
<b>Emotions</b>	2	Enjoyment, happiness
	1	Engagement, focus
	0	Neutral
	-1	Anxiety, frustration
	-2	Boredom, laziness
<b>Attitude</b>	-3	Fear, displeasure
	1	Enthusiastic, energetic
	0	Proper
	-1	Lazy
<b>Gaze</b>	-2	Do not train
	1	Look at the robot
	0	Look at himself
	-1	Look at others
<b>Communication</b>	-2	Not involved
	2	Speak and gestures
	1	Speak or gestures
	0	Hear the robot
	-1	Speak to others

The interaction level is different throughout the session, so we thought it convenient to divide the sessions into 6 logical segments to analyze the child’s reactions separately. Using continuous data from the video annotations, we calculate the Interaction Level (IL) metric to find the quality of the interaction for each segment. To obtain the IL, we calculate the average duration for each behavior of each annotation track and then normalize these durations by dividing them by the average of the total duration of the segment. Next, we multiply the values calculated for each behavior by the corresponding score shown in Table 1. Finally we add all behavior values together for every aspect of interaction (Emotions, Gaze, Communication and Attitude) and apply Equation 2, which is an adaptation of Fridin’s

work [16] to use continuous duration values. Communication and attitude are more relevant than the other aspects in achieving a successful interaction, so their contribution to the final IL value is doubled. In our case, the minimum value is -11 and the maximum is +9. We do these calculations for each segment and for the whole session which is considered as an individual segment.

$$IL = Emotions + Gaze + 2(Commun. + Attitude) \quad (2)$$

We also evaluate each pose with an adaptation of the performance metric proposed by Fridin [16]. Its value is 3 if the children carry out the movement correctly at the first attempt, 2 at the second attempt, 1 at the third attempt and 0 if he cannot carry out the pose at all.

## 7 Evaluation of the Child-Robot Interaction

NAOTherapist has been evaluated using more than one hundred healthy children in schools using short therapy sessions and with three real patients using full-length sessions. We have used a large number of questionnaires and video data to evaluate the child-robot interaction with the developed architecture. For this evaluation, the robotic platform follows the use case for every participant.

Table 2 shows the average features of the executed sessions for the 117 healthy children from two schools and 3 pediatric patients. These results include different average calculations of the sessions evaluated: the duration of sessions, the number of planning actions executed by the robot (including exogenous events to finish the session) and percentage of possible attempts made, corrections and skipped or omitted poses. When calculating these results, attempts are considered since the first execution of the pose until the last required correction. This means that a participant always has at least one attempt. Corrections depend on the success of the poses made. So the minimum number of attempts is the number of poses in the session (1 each) and the maximum is the product of the number of poses from the three possible attempts.

As can be seen in Table 2, the sessions at the hospital comprise a higher number of poses than at the school. Furthermore, the patients used the 61% of the possible attempts, opposed to the healthy children who only needed 24%.

### 7.1 Schoolchildren’s Questionnaires

Table 2 also shows the results of the questionnaires. A result below 0.5 is undesirable for us, but we highlight

answers below 0.7 to clarify those that have the worst results. Questions were coded from Q1 to Q19b, in a useful order for us. the results of Q9, Q16 and Q17 are just informative and we do not have any particular preference.

**Table 2** Features and Questionnaires of the Evaluations

Participants	Schools	Hospital		
	117	A	B	C
Condition	Healthy	OBPP	OBPP	CP
Age	7.90 ± 1.4	7	9	7
Gender (0=M, 1=F)	0.45	0	0	0
Duration (s)	296 ± 50	772	912	831
Num. actions	65.82 ± 4.6	140	148	146
Min-Max attempts	21.7 – 65.2	44 – 132		
Needed attemp. (%)	24.18 ± 6.7	57.6	63.6	62.1
Corrections (%)	16.12 ± 6.8	36.4	45.5	43.2
Failed poses (%)	9.65 ± 7.0	22.7	31.8	27.3
Q1	0.87 ± 0.3	1	1	1
Q2	0.58 ± 0.5	0	0	0
Q3	0.88 ± 0.2	1	1	1
Q4	0.91 ± 0.3	1	0	1
Q5	0.68 ± 0.5	0	0	0
Q6	0.67 ± 0.3	0.5	0	0
Q9	6.86 ± 4.3	0	10	6
Q10	0.98 ± 0.1	1	0	1
Q11	0.94 ± 0.2	1	0.5	1
Q12	0.87 ± 0.3	1	1	1
Q13a	0.95 ± 0.2	1	1	1
Q13b	0.84 ± 0.3	0	1	1
Q13c	0.97 ± 0.1	1	1	1
Q13d	1.00 ± 0.0	1	1	1
Q15	0.39 ± 0.5		–	
Q16	0.48 ± 0.5		–	
Q17	0.74 ± 0.4		–	
Q18a	0.92 ± 0.2	1	1	1
Q18b	0.81 ± 0.4	1	0	1
Q18c	0.88 ± 0.3	1	1	1
Q19a	0.95 ± 0.1	1	1	1

Almost all schoolchildren decided that it was easy to understand what they had to do with the robot (Q1). There are many differences between the children when they had to decide if the robot was alive or not (Q2). All the children felt that the robot was gazing at them (Q3) but they were not overwhelmed by it (Q4). There are more differences when they have to evaluate whether the robot spoke too much (Q5). We observed that some children wanted to have a physical interaction with the robot, or that they were tired of hearing corrections when they were repeatedly doing the exercises wrong. The question about whether the robot had feelings or not (Q6) has similar results to Q2. When the children had to guess the age of the robot (Q9), we observed

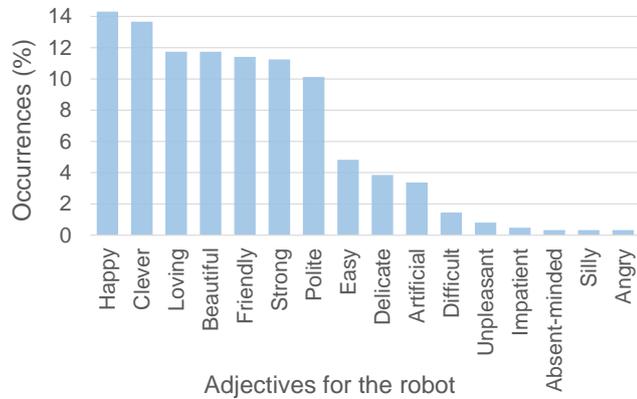
that they thought that the robot was a little younger than them. Almost all the schoolchildren agreed that they wanted to have the robot at home (Q10) and even to be attended by it in the hospital (Q11). Q11 has the opposite result than in the previous work of Therapist [5]. This may be because the NAO robot is smaller than the children, which could make it less intimidating and friendlier than the Ursus robot used in the Therapist project. Furthermore, children did not think that they were scolded by the robot (Q12). They thought that the robot could see them (Q13a) and, surprisingly, also hear them (Q13b), although our system does not have audio recognition capabilities yet. All participants thought that the robot enjoyed playing with them (Q13c) and, if they had to do physiotherapy in hospital, they would rather do it with the robot (Q13d).

The question about whether the robot was correcting a pose which indeed was correct (Q15), had an undesirable result, although the children had problems understanding this question. The system rarely fails when correcting poses, but many children could not understand that they had to put their arms in exactly the same position as the robot showed them. Moreover, even with the eyes changing dynamically from red to green according to the correctness of the pose, some children found it difficult to coordinate their own arms when making the exact pose. The lack of a mirror in front of the participant makes this task difficult, but coordination in this imitation activity is important for the success of the physiotherapy.

Both exercises looked the same (Q16) and the second one was considered more difficult (Q17), as was intended. They also consider that the descriptions of the exercises were easy to understand (Q18a) and the session was not exhausting (Q18b). The feedback with the lights of the eyes, as described in Section 4, was useful (Q18c). Finally, children do not think that the session was boring (Q19a).

Participants also had to select about 5 adjectives from a list of 16 (Q7), as in the previous work of Therapist [5]. Figure 6 presents the list of all adjectives with the proportion of the selected ones. Clearly, all adjectives with a positive connotation have been selected in the first place, which is evidence of the children’s acceptance of the system (hypothesis H2). Some of these adjectives like “easy” are used for artificial entities instead of social ones. Each adjective has a positive or negative value according to its connotation and application to a social entity as explained in Section 6.3. The social vs. artificial metric is calculated by adding all these values together for each child. The average of this metric for each child is 2.475, which indicates that

the robot was mostly considered as a social entity validating hypothesis H3.



**Fig. 6** Proportion of adjectives selected by the children to describe the robot (Q7).

The children also had to give the robot a name (Q8). This question is difficult to evaluate, but teachers and family confirmed that they often tend to put their own name, a friend’s or their pet’s name. Older children were more creative with fictitious names. We also asked for more games they would like to play with the robot (Q14). The majority of them involved physical activities like playing with a ball, running, etc. This suggests that children love to see the robot moving by itself. The final question was free; about whether they liked playing with the robot or not (Q19b). The majority said that they had a lot of fun with the robot because of the way it moves and speaks. Some of them said that they would like to see the robot walking, moving its legs and to be closer to touch it. This question was useful to see the children’s expectations for future improvements in the system.

In conclusion, we can confirm that schoolchildren did not have any problem following the sessions. They mostly considered the robot as a social entity, although not necessarily alive. The results of the questionnaire show a huge acceptance of the robotic system in all evaluations, as a playmate and as a tool to support their physical rehabilitation. These results are consistent with hypotheses H2 and H3.

## 7.2 Video Data Analysis

We carried out in-depth analysis of the videos of the last 50 schoolchildren because they shared the same set of poses and were very comparable between them. The duration of the session is divided into 6 logical segments, containing different activities. In the first-contact segment, the robot wakes up, says “hello” and introduces

itself. Then, in the introduction, the robot explains the task that they are going to do to the child. Then, they do a warm-up exercise and a dissociation exercise. Finally, the robot says “good-bye” and, in the parting segment, it sits down and goes to sleep again. Almost the 80% of the time of the session is spent doing exercises and the rest is social interaction with the robot. Our metrics on the video data are based on continuous time values, so we think that it is important to consider each segment of the session individually to extract conclusions from the analysis. All of these metrics were explained in Section 6.3.

Table 3 summarizes the results of the analysis of the annotations for each segment and the full session considered as an individual segment. 5 different types of annotation, or aspects, are shown in this table (E: Emotions, A: Attitude, G: Gaze, C: Communication). The sum of the percentages is 100% for each behavior and each segment. In general, the standard deviations are high, but we can extract several conclusions in some segments and behaviors. The parting segment has the worst results because children often do not wait for the robot until it is fully seated. They did this to avoid delaying the next participant and start the questionnaire quickly. Annotations on emotions show that most of the time the child is just focused on performing doing the exercises correctly. Children spend more time enjoying segments which are not exercises because they require social interaction. Displeasure values are produced mostly in parting because sometimes children left the robot before it finished the sitting down animation. In the annotation of attitude, we consider that for the majority of the time the children are well behaved. This is followed by the enthusiastic behavior, corresponding to very motivated children. Almost none of children were apathetic with the robot and during the training session all of them followed the instructions completely. These results are consistent with hypotheses H1, H2 and H3.

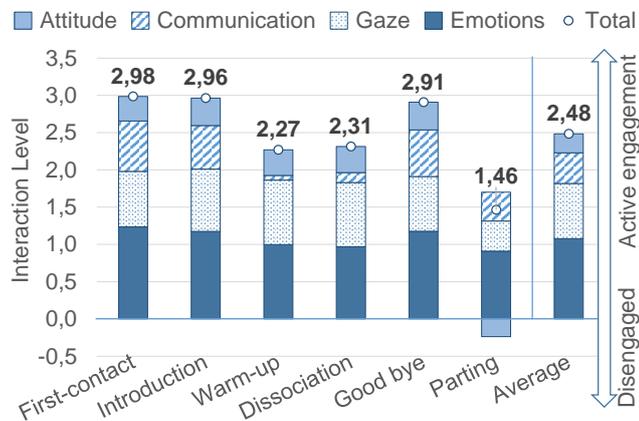
Almost all the time children were gazing at the robot. Children rarely look themselves to check their posture and, more frequently, they look away to the observers or other children in the experimental room looking for some kind of feedback. Children usually respond verbally (sometimes shyly) to the robot when it says “hello”, “good-bye” and asks how they are. These communications are short but very valuable because they imply an active social interaction (hypothesis H3).

A graphical view of the interaction is shown in Figure 7. This figure shows the interaction level metric for each segment and the contribution for each aspect of interaction. Higher levels of interaction are reached in segments in which there are no exercises, because these

**Table 3** Behavior Distribution throughout the Segments of a Session

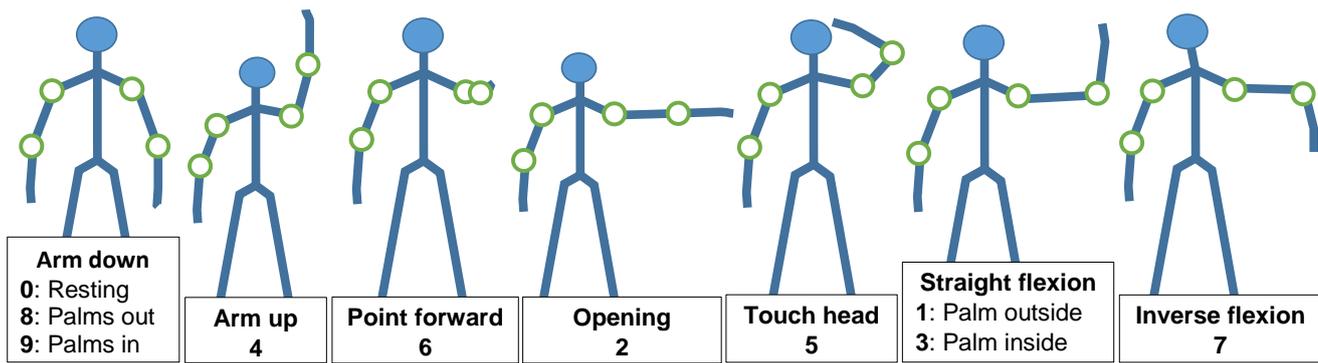
Behavior (%)	First contact	Introduction	Warm-up	Dissociation	Good bye	Parting	Full session
<b>E - Enjoyment</b>	44.09	28.48	7.94	13.70	30.72	26.42	16.31 ± 19.3
<b>E - Engagement</b>	39.24	60.48	84.59	72.11	62.37	46.29	71.48 ± 26.1
<b>E - Neutral</b>	15.35	11.04	6.87	11.94	4.69	24.18	10.60 ± 18.5
<b>E - Frustration</b>	0.00	0.00	0.60	1.83	0.00	0.00	1.03 ± 2.4
<b>E - Boredom</b>	0.00	0.00	0.00	0.42	0.59	1.21	0.29 ± 1.4
<b>E - Displeasure</b>	1.31	0.00	0.00	0.00	1.64	1.90	0.28 ± 1.4
<b>A - Enthusiastic</b>	19.69	23.04	21.52	19.89	20.52	19.34	20.56 ± 30.9
<b>A - Proper</b>	79.00	72.96	74.51	79.00	75.62	64.42	76.38 ± 34.6
<b>A - Lazy</b>	0.00	4.48	4.23	2.26	2.23	2.25	2.84 ± 9.1
<b>A - Do not play</b>	1.57	0.00	0.00	0.00	0.00	14.51	0.78 ± 1.2
<b>G - Look robot</b>	87.40	91.36	92.58	93.00	88.04	76.86	91.34 ± 17.6
<b>G - Look himself</b>	0.00	0.00	0.99	1.45	0.00	0.00	0.97 ± 1.8
<b>G - Look others</b>	11.15	8.16	6.32	6.26	13.48	10.71	7.39 ± 9.3
<b>G - Distracted</b>	1.05	0.00	0.00	0.00	0.00	12.78	0.67 ± 1.1
<b>C - Voice + gestures</b>	14.57	8.48	1.65	3.54	12.31	13.82	4.98 ± 9.0
<b>C - Voice / gestures</b>	8.40	12.32	0.44	0.95	7.15	7.43	2.56 ± 2.4
<b>C - Hear robot</b>	72.05	78.08	97.73	95.05	81.36	63.04	91.14 ± 21.0
<b>C - Speak others</b>	4.07	0.32	0.57	1.25	0.23	15.72	1.78 ± 2.5

segments are only based on social interaction. Emotions and communication are clearly lower in segments with exercises because focusing on training is enough to do them correctly. Attitude and gaze are the same in all segments (except in parting) as the child is almost always looking at the robot to follow its instructions. In parting, attitude has a negative contribution because children do not wait until the robot is fully seated. All segments show an active engagement of the children. This is consistent with hypotheses H1, H2 and H3.

**Fig. 7** Average Interaction Level (IL) distribution throughout the segments of the session.

In these experiments, the postures of the arms are intended to be imitated easily by healthy children. Moreover, we wanted to test a hard, unnatural posture for them to give rise to a lot of corrections. This posture requires the elbow to be maintained at the shoulder height and the hand down at an angle of 90 degrees to the elbow joint. This is identified with a 7 in our system (inverse flexion), as shown in Figure 8. The resting posture has the identifier 0 and it is not considered when comparing the pose. Postures 8 and 9 and postures 1 and 3 differ only in wrist rotations. These differences cannot be detected accurately with the skeleton-tracking algorithm of Windows Kinect SDK, so they are compared as the same pose.

Figure 9 shows a bar for every pose in the sessions in order with the average value of the performance metric. The name of the pose contains the code of the posture for each arm. Poses with the posture 7 (the unnatural one) have low performance, as we expected. Postures 8 and 9 only require the arms to be down with different wrist angles, so their performance value is high. The last pose (6-6) is simple, but confusing in practice. In this one, both arms must be straight and pointing out in front. The children usually believed that they had to point at the robot with their arms, lowering them too much because the NAO robot is shorter than them. Sometimes this pose is well done, but the Vision component has problems in comparing the angles of the joints



**Fig. 8** Frontal diagrams and numeric identifiers for each tested posture in our system. In this figure, the right arm has always the posture 0.

because the arms are perpendicular to the plane of the Kinect sensor.

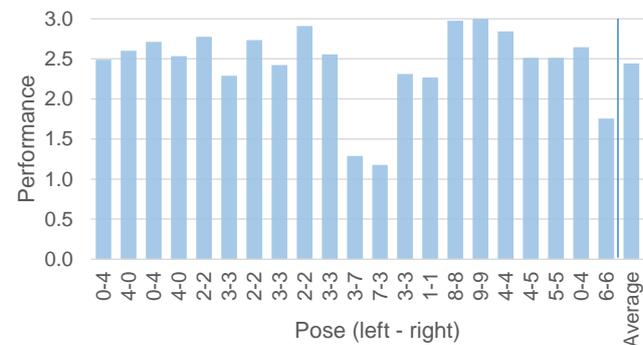
The first poses of the session contains posture 4, which requires the arms to be straight and up. In these first poses, the children tend to raise their arms shyly, with their hands at the height of the head. Similar problems are found in posture 3 (the same as in 7, but with the hands up). After the first corrections, the children get the clue from the color of the eyes and they know how to do the exercises much better for the following poses (hypothesis H4). We observed small detection problems in posture 4 when children have thin complexion, wearing a scarf or have long hair in front of their shoulders. In all cases the session was able to continue normally. The children smile with posture 5, which requires a hand on top of the head.

The results of the analysis of the video annotations are coherent with the observers' comments and the questionnaires. The children were focused on the activity, they enjoyed the session trying to do the exercises as well as possible and they interacted socially with the robot. The robot is able to do the full session autonomously with no problems. Therefore, video data support hypotheses H1 to H5.

## 8 Evaluation With Pediatric Patients

The last evaluation was carried out with 3 males<sup>2</sup>, two seven year-olds and one nine year-old. They are pediatric patients from the *Hospital Universitario Virgen del Rocío* (HUVR). Two of them have obstetric brachial plexus palsy (OBPP) and the other suffers from cerebral palsy (CP). In some cases, they exhibit some degree of dystonia (twisting and unintentional

movements) while performing the exercises. The experimental conditions were very similar to the previous evaluations. 4 exercises were used instead of 2: warming up, maintaining poses, dissociation poses and cooling down. Each child had his own motor disabilities, but the exercises in all of the sessions were the same for experimental purposes. The experimental room chosen was where these children usually do their physiotherapy exercises. However, in this case, there were observers such as physicians, therapists and technicians who, after the session, also filled in a different questionnaire. Next to the training area, there was a window from which the child's family and other observers were able to watch the therapy session.



**Fig. 9** Performance measurements for each pose. A 0 means that the child failed to make the pose after three attempts, and a 3 means that the children performed the pose at the first try. Each pose contains the code of the posture for the left and right arm, separated by a hyphen.

The children did the exercises well, in spite of them lasting about 15-20 minutes of rehabilitation, which for them is long. The children were used to do similar rehabilitation movements and they understood the procedure quickly. The dynamic-comparison threshold was more permissive when the child failed several consecu-

<sup>2</sup> Online videos of the evaluations in the HUVR:

Patient A: <https://youtu.be/9n9n1128rME>

Patient B: <https://youtu.be/77a20MzLVwQ>

Patient C: [https://youtu.be/kV-\\_b-sd54I](https://youtu.be/kV-_b-sd54I)

tive times. This avoided too many corrections for the same child.

The questionnaires for children (Table 2) were the same as those for the school, although the questions had to be explained by adults. Questions which required writing (Q7, Q8, Q14 and Q19b) or evaluating technical aspects of the exercises (Q15, Q16 and Q17) were not answered by all participants, so they were not assessed. The results have several interesting differences from those from the school, although pediatric patients are too few to be representative enough. They thought that the robot was not alive (Q2), but it had “some feelings” (Q6). All of them thought that the robot spoke too much (Q5), probably because it was the first time that we tested the system with full-length sessions and they had to make many corrections, in spite of all of them agreeing that the session was fun and productive (Q19a). The children considered the robot a therapeutic toy because they all agreed to do more physiotherapy sessions with it (Q13d).

There were different duration requirements when designing the sessions for schoolchildren and pediatric patients. The sessions in schools lasted about 5 minutes while in the hospital reached 15 minutes. This difference gave patients more time to realize that the robot was not able to hear them (Q13b) and they found the session more tiring (Q18). The latter could be the reason why one patient would rather not have the robot at home (Q10).

The physicians and the therapists thought that the robot was a very useful tool. A physician detected certain clinical aspects on a participant that she never realized before. The children were uninhibited with the robot and, when repeating and performing movements, some unseen limitations or capacities could have appeared. So the robotic system has proven to be a useful tool for diagnosis too.

After each patient’s session, the respective family, two physicians and a therapist filled in a questionnaire whose results are shown in Table 4. As a reminder, the answers to the questionnaires are represented from 0 to 1, 1 being the most positive result in our evaluations. All questions obtained very positive results although there are some differences between each group. Both the family and the therapists thought that the children had understood what to do (Q1), but sometimes the physicians did not think so. In general, the movements of the robot are natural (Q2), the children carried out all poses naturally (Q3) and they were not overwhelmed with the session (Q4). For therapists, Q2, Q3 and Q4 did not produced the most desirable answer because, for evaluation purposes, all exercises were the same in all sessions and, consequently, they were not adapted to

the child’s requirements. All observers agreed on all the following questions: the robot only corrected incorrect poses (Q5), the sessions were carried out by the robot fluently (Q6), the children were engaged in the session (Q7), this was a beneficial experience for them (Q8), the patients made an effort to do the exercises (Q9) and finally that the robot was a useful tool in rehabilitating children with these medical conditions (Q10). These results reinforce hypothesis H6, although to establish the final conclusions, a wider, long-term evaluation with more pediatric patients is required [26].

**Table 4** The results of the Questionnaires for Observers and Experts

	Family	Physicians	Therapists	Total
<b>Q1</b>	1.00	0.67	0.83	0.79 ± 0.3
<b>Q2</b>	1.00	1.00	0.50	0.88 ± 0.2
<b>Q3</b>	1.00	0.92	0.67	0.88 ± 0.2
<b>Q4</b>	1.00	0.75	0.42	0.73 ± 0.3
<b>Q5</b>	1.00	1.00	1.00	1.00 ± 0.0
<b>Q6</b>	1.00	1.00	1.00	1.00 ± 0.0
<b>Q7</b>	1.00	0.92	1.00	0.96 ± 0.1
<b>Q8</b>	1.00	1.00	1.00	1.00 ± 0.0
<b>Q9</b>	0.83	0.92	1.00	0.92 ± 0.2
<b>Q10</b>	1.00	1.00	1.00	1.00 ± 0.0

## 9 Conclusion

The evaluation presented in this work has been carried out with more than 120 children. Our architecture is able to perform all physiotherapy sessions autonomously without the need for human intervention (H5). Although the results of the questionnaires reveal that not all participants consider that the robot was alive, the behavior, speech and appearance of the robot guarantee its social prominence in spite of the fact that there were always other observers in the room (H3).

According to the results of the interaction, the participants enjoyed themselves while training with the NAO robot (H2) and they have shown themselves to be motivated and engaged (H1). In fact, there were children who had more difficulties achieving certain poses, but they did not give up trying to surpass themselves. In most cases; the children figured out how to train with the robot without any help (H4) and, after few attempts and corrections, they managed to perform the rest of the exercises correctly by themselves. The videos of the pediatric patients show the great effort made by them during the physiotherapy session. When playing with a

robot, children become uninhibited, having an active engagement and being committed to the exercises.

Our experiments involve only one session for each child, always having their first contact with the robot. The results are very promising because children want to repeat the experience, but it would be necessary to carry out long-term experiments to decide whether the children's engagement is maintained over time (H6). Experts have an optimistic attitude in this regard. Few children currently have the opportunity to interact with a social robot like NAO, so the chance to play with it gives an interesting plus to the physiotherapy therapy. The children could find new motivation to continue their treatment by playing with the robot.

The deployment of the NAOTherapist platform is agile and not very expensive, so it seems to be an interesting investment for a hospital or a children's physiotherapy center. Our system may be considered as a novel physiotherapy service assisted by a humanoid robot whose beneficiaries are not only patients but also physicians and therapists, since our system could be a new objective tool for diagnosis.

Moreover, the NAOTherapist architecture is one of the few whose execution of the rehabilitation therapy is carried out autonomously and has already had a warm reception from the children, their family and experts. Its later integration into the Therapist project will allow the incorporation of more functions such as clinical metrics capture, clinical reports generation, facial recognition or voice interaction.

Our new challenges should focus on the capability of the robot to change and maintain their empathy with the patient throughout all of the sessions of his therapy. In this sense, the robot should provide new behaviors and games which the patient may consider attractive to play and maintain or increase adherence to the physiotherapy treatment.

## A Children's Questionnaire

- Q1. Was it easy to understand what to do with the robot?
- Q2. Do you think the robot is alive?
- Q3. Do you think the robot was gazing at you?
- Q4. Did you feel overwhelmed when the robot talked to you?
- Q5. Do you think the robot speaks too much?
- Q6. Do you think the robot has feelings?
- Q7. Choose 5 adjectives to describe the robot
- Q8. What name would you give to the robot?
- Q9. How old do you think the robot is?
- Q10. Would you like to have this robot at home?
- Q11. Would you like to be treated by the robot?
- Q12. Do you think the robot can see you?
- Q13a. Do you think the robot can hear you?
- Q13b. Do you think the robot is glad when you play together?
- Q13c. Would you like to do more exercises with the robot?
- Q13d. Which games would you want to play with the robot?

- Q15. Did the robot correct an actual correct pose?
- Q16. Which exercise did you like most?
- Q17. Which exercise was the most difficult?
- Q18a. Did you understand the descriptions of the exercises?
- Q18b. Were the exercises tiring?
- Q18c. Did the lights of the eyes help you to do the exercises?
- Q19a. Were the exercises boring?
- Q19b. Why?

## B Observers and Experts' Questionnaire

- Q1. Did the child understand what to do?
- Q2. Are the movements of the robot natural?
- Q3. Did the child perform the movements naturally?
- Q4. Was the child overwhelmed during the session?
- Q5. Did the robot correct an actual correct pose?
- Q6. Was the session carried out fluently?
- Q7. Was the child very committed to the session?
- Q8. Was this experience beneficial for the child?
- Q9. Did the child make a great effort to finish the session?
- Q10. Is this system a useful tool for physiotherapy?

## References

1. Alcázar V, Guzmán C, Prior D, Borrajo D, Castillo L, Onaindia E (2010) PELEA: Planning, Learning and Execution Architecture. In: Proceedings of the 28th Workshop of the UK Planning and Scheduling Special Interest Group (PlanSIG)
2. Boccanfuso L, O'Kane JM (2011) Charlie : An adaptive robot design with hand and face tracking for use in autism therapy. *International Journal of Social Robotics* 3(4):337–347, DOI 10.1007/s12369-011-0110-2
3. Borggraefe I, Kiwull L, Schaefer JS, Koerte I, Blaschek a, Meyer-Heim a, Heinen F (2010) Sustainability of motor performance after robotic-assisted treadmill therapy in children: an open, non-randomized baseline-treatment study. *European journal of physical and rehabilitation medicine* 46(2):125–31
4. Burgar CG, Lum PS, Shor PC, Van der Loos HM (2000) Development of robots for rehabilitation therapy: the Palo Alto VA/Stanford experience. *Journal of rehabilitation research and development* 37(6):663–674
5. Calderita VL, Manso JL, Bustos P, Suárez-Mejías C, Fernández F, Bandera A (2014) THERAPIST: Towards an Autonomous Socially Interactive Robot for Motor and Neurorehabilitation Therapies for Children. *JMIR Rehabilitation and Assistive Technologies (JRAT)* 1(1):e1, DOI 10.2196/rehab.3151
6. Castelli E (2011) Robotic movement therapy in cerebral palsy. *Developmental Medicine & Child Neurology* 53(6):481–481, DOI 10.1111/j.1469-8749.2011.03987.x
7. Choe Yk, Jung HT, Baird J, Grupen RA (2013) Multidisciplinary stroke rehabilitation delivered by a humanoid robot: Interaction between speech and physical therapies. *Aphasiology* 27(3):252–270, DOI 10.1080/02687038.2012.706798
8. Dehkordi PS, Moradi H, Mahmoudi M, Pouretemad HR (2015) The design, development, and deployment of roboparrot for screening autistic children. *International Journal of Social Robotics* 7(4):513–522, DOI 10.1007/s12369-015-0309-8

9. Drubicki M, Rusek W, Snela S, Dudek J, Szczepanik M, Zak E, Durmala J, Czernuszenko A, Bonikowski M, Sobota G (2013) Functional effects of robotic-assisted locomotor treadmill therapy in children with cerebral palsy. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine* 45(4):358–63, DOI 10.2340/16501977-1114
10. Dubowsky S, Genot F, Godding S, Kozono H, Skwersky A, Yu H, Yu LS (2000) Pamm-a robotic aid to the elderly for mobility assistance and monitoring: a helping-hand for the elderly. In: *Robotics and Automation, 2000. Proceedings. ICRA'00. IEEE International Conference on, IEEE, vol 1*, pp 570–576
11. Eriksson J, Mataric MJ, Winstein C (2005) Hands-off Assistive Robotics for Post-Stroke Arm Rehabilitation. In: *Proceedings of the 9th International Conference on Rehabilitation Robotics (ICORR)*, IEEE, pp 21–24
12. Fasola J, Mataric M (2010) Robot exercise instructor: A socially assistive robot system to monitor and encourage physical exercise for the elderly. In: *RO-MAN, 2010 IEEE*, pp 416–421, DOI 10.1109/ROMAN.2010.5598658
13. Feil-Seifer D, Mataric MJ (2005) Defining Socially Assistive Robotics. In: *Proceedings of the 9th International Conference on Rehabilitation Robotics (ICORR)*, IEEE, pp 465–468
14. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. *Robotics and autonomous systems* 42(3):143–166
15. Fox M, Long D (2003) PDDL2.1: An Extension to PDDL for Expressing Temporal Planning Domains. *Journal of Artificial Intelligence Research (JAIR)* 20(1):61–124
16. Fridin M (2014) Kindergarten social assistive robot: First meeting and ethical issues. *Computers in Human Behavior* 30(0):262 – 272, DOI <http://dx.doi.org/10.1016/j.chb.2013.09.005>
17. Fridin M, Belokopytov M (2014) Robotics agent coacher for cp motor function (rac cp fun). *Robotica* 32:1265–1279, DOI 10.1017/S026357471400174X
18. Garcia N, Sabater-Navarro J, Gugliemeli E, Casals A (2011) Trends in rehabilitation robotics. *Medical & Biological Engineering & Computing* 49(10):1089–1091, DOI 10.1007/s11517-011-0836-x
19. Ghallab M, Nau D, Traverso P (2004) *Automated Planning: Theory & Practice*. Elsevier
20. Gonzalez JC, Pulido JC, Fernandez F (2016) A three-layer planning architecture for the autonomous control of rehabilitation therapies based on social robots. *Cognitive Systems Research*, DOI 10.1016/j.cogsys.2016.09.003
21. Graf B, Reiser U, Hägele M, Mauz K, Klein P (2009) Robotic home assistant care-o-bot®3 - product vision and innovation platform. In: *Advanced Robotics and its Social Impacts (ARSO), 2009 IEEE Workshop on*, pp 139–144, DOI 10.1109/ARSO.2009.5587059
22. Hoffmann J (2003) The Metric-FF Planning System: Translating “Ignoring Delete Lists” to Numeric State Variables. *Journal of Artificial Intelligence Research (JAIR)* 20(1):291–341
23. Kahn LE, Averbuch M, Rymer WZ, Reinkensmeyer DJ, D P (2001) Comparison of robot-assisted reaching to free reaching in promoting recovery from chronic stroke. In: *Integration of Assistive Technology in the Information Age, Proceedings 7th International Conference on Rehabilitation Robotics*, IOS Press, pp 39–44
24. Kozima H, Michalowski MP, Nakagawa C (2008) Keepon. *International Journal of Social Robotics* 1(1):3–18, DOI 10.1007/s12369-008-0009-8
25. Lacey G, Dawson-Howe KM (1998) The application of robotics to a mobility aid for the elderly blind. *Robotics and Autonomous Systems* 23(4):245 – 252, DOI [http://dx.doi.org/10.1016/S0921-8890\(98\)00011-6](http://dx.doi.org/10.1016/S0921-8890(98)00011-6), intelligent Robotics Systems - SIRS'97
26. Leite I, Martinho C, Paiva A (2013) Social robots for long-term interaction: A survey. *International Journal of Social Robotics* 5(2):291–308, DOI 10.1007/s12369-013-0178-y
27. Manso L, Bachiller P, Bustos P, Núñez P, Cintas R, Calderita L (2010) RoboComp: A Tool-Based Robotics Framework. In: Ando N, Balakirsky S, Hemker T, Reggiani M, von Stryk O (eds) *Simulation, Modeling, and Programming for Autonomous Robots, Lecture Notes in Computer Science*, vol 6472, Springer Berlin Heidelberg, pp 251–262, DOI 10.1007/978-3-642-17319-6\_25
28. Manso LJ, Calderita LV, Bustos P, García J, Martínez M, Fernández F, Garcés AR, Bandera A (2014) A general-purpose architecture to control mobile robots. In: *Proceedings of the 15th Workshop of physical agents (WAF 2014)*, León, Spain, pp 105–116
29. Mataric M, Eriksson J, Feil-Seifer D, Winstein C (2007) Socially assistive robotics for post-stroke rehabilitation. *Journal of NeuroEngineering and Rehabilitation* 4(1):5, DOI 10.1186/1743-0003-4-5
30. McMurrough C, Ferdous S, Papangelis A, Boisselle A, Heracleia FM (2012) A survey of assistive devices for cerebral palsy patients. In: *Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments, ACM, New York, NY, USA, PETRA '12*, pp 17:1–17:8, DOI 10.1145/2413097.2413119
31. Meyer-Heim A, van Hedel HJ (2013) Robot-assisted and computer-enhanced therapies for children with cerebral palsy: Current state and clinical implementation. *Seminars in Pediatric Neurology* 20(2):139 – 145, DOI <http://dx.doi.org/10.1016/j.spen.2013.06.006>, update on Cerebral Palsy: Diagnostics, Therapies and the Ethics of it All
32. Nalin M, Baroni I, Sanna A (2012) *A Motivational Robot Companion for Children in Therapeutic Setting*. In: *IROS 2012*
33. Nau D, Au TC, Ilghami O, Kuter U, Murdock JW, Wu D, Yaman F (2003) SHOP2: An HTN Planning System. *Journal of Artificial Intelligence Research (JAIR)* 20:379–404
34. Ni D, Song A, Tian L, Xu X, Chen D (2015) A walking assistant robotic system for the visually impaired based on computer vision and tactile perception. *International Journal of Social Robotics* 7(5):617–628, DOI 10.1007/s12369-015-0313-z
35. Perry J, Rosen J, Burns S (2007) Upper-limb powered exoskeleton design. *Mechatronics, IEEE/ASME Transactions on* 12(4):408–417, DOI 10.1109/TMECH.2007.901934
36. Pulido JC, González JC, González-Ferrer A, García J, Fernández F, Bandera A, Bustos P, Suárez C (2014) Goal-directed Generation of Exercise Sets for Upper-Limb Rehabilitation. In: *Proceedings of Knowledge Engineering for Planning and Scheduling workshop (KEPS), ICAPS*, pp 38–45
37. Song A, Wu C, Ni D, Li H, Qin H (2016) One-therapist to three-patient telerehabilitation robot system for the upper limb after stroke. *International Journal of Social Robotics* 8(2):319–329, DOI 10.1007/s12369-016-0343-1

38. Suárez Mejías C, Echevarría C, Núñez P, Manso L, Bustos P, Leal S, Parra C (2013) Ursus: A Robotic Assistant for Training of Children with Motor Impairments. In: *Converging Clinical and Engineering Research on Neurorehabilitation, Biosystems & Biorobotics*, vol 1, Springer Berlin Heidelberg, pp 249–253, DOI 10.1007/978-3-642-34546-3\_39
39. Tapus A, Mataric M, Scasselati B (2007) Socially assistive robotics [Grand Challenges of Robotics]. *Robotics Automation Magazine*, IEEE 14(1):35–42, DOI 10.1109/MRA.2007.339605
40. Wainer J, Dautenhahn K, Robins B, Amirabdollahian F (2013) A pilot study with a novel setup for collaborative play of the humanoid robot kaspar with children with autism. *International Journal of Social Robotics* 6(1):45–65, DOI 10.1007/s12369-013-0195-x