



**IM-TWIN: from Intrinsic Motivations  
to Transitional Wearable INTelligent  
companions for autism spectrum disorder**  
*a European funded project*

***Empirical validation: PlusMe***  
**Deliverable 4.1**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952095.

Project duration 36 months (November 2020, October 2023).  
Consortium: Consiglio Nazionale delle Ricerche (ITA),  
Universiteit Utrecht (NLD), Centre de Recherches  
Interdisciplinaires (FRA), Università degli Studi di Roma  
La Sapienza (ITA), Plux-Wireless Biosignals S.A. (PRT).

## Deliverable data

<b>Work Package:</b>	4 Validation of the PlusMe and IM-TWIN system
<b>Work Package leader:</b>	LA SAPIENZA
<b>Deliverable beneficiary:</b>	LA SAPIENZA, CRI, CNR
<b>Dissemination level:</b>	Public
<b>Due date:</b>	31 <sup>th</sup> October (Month 24)
<b>Type:</b>	Report
<b>Revision:</b>	2 (March 2023)
<b>Authors:</b>	F. Giocondo, L. Jacquy, N. Faedda, G. Cavalli, C. Sogos, J.K. O'Regan, V. Guidetti, V. Sperati, G. Baldassarre

## Acronyms of partners

CNR-ISTC	Consiglio Nazionale delle Ricerche, Istituto di Scienze e Tecnologie della Cognizione (Italy)
UU	Universiteit Utrecht (The Netherlands)
CRI	Centre de Recherches Interdisciplinaires (France)
LA SAPIENZA	Università degli Studi di Roma La Sapienza (Italy)
PLUX	Plux - Wireless Biosignals S.A. (Portugal)

# Table of contents

<b>1. Overview of the deliverable</b>	<b>3</b>
<b>2. LA SAPIENZA: PlusMe as a diagnostic tool</b>	<b>3</b>
2.1. Aim of the experiment	3
2.2. Participants	4
2.3. Procedure	4
2.3.1. Experimental setting	4
2.3.2. Experimental protocol	5
2.4. Data collection	6
2.5. Preliminary results	7
<b>3. LA SAPIENZA: PlusMe as a therapeutic tool</b>	<b>8</b>
3.1. Aim of the experiment	8
3.2. Participants	9
3.3. Procedure	11
3.4. Data collection	12
3.5. Results and discussion	12
<b>4. CRI: PlusMe and the early diagnosis of ASD</b>	<b>13</b>
4.1. Aim of the experiment	13
4.2. Participants	14
4.3. Procedure	15
4.3.1. Experimental setting	15
4.3.2. Experimental protocol	16
4.4. Data collection	16
4.5. Results and discussion	17
<b>5. Future Developments</b>	<b>19</b>
5.1. New experimentation and testing	20
5.1.1. PlusMe and App	20
5.1.1.2. Experiments with new PlusMe (CRI)	21
5.1.2. Camera glasses	21
5.1.3. Octopus X-8 device	22
5.1.4. Experiment with sensorised t-shirt	23

# 1. Overview of the deliverable

The deliverable reports some results of experiments conducted in France by CRI and in Italy by LA SAPIENZA. The experiment in France involved 38 Typically Developed (TD) children. The aim was to analyse the age at which *PlusMe* is useful as a tool to make an early diagnosis of ASD. During the experiments in Italy the *PlusMe* was used in two main fields:

- In the therapy of ASD children, to validate the toy as a therapeutic tool to increase some social behaviours;
- In the early diagnosis of ASD, to validate the toy as a diagnostic tool to detect warning signals of autism.

## 2. LA SAPIENZA: *PlusMe* as a diagnostic tool

### 2.1. Aim of the experiment

The *PlusMe* device is used to assess its potential use as a screening tool for signal risk of Autism Spectrum Disorder. The experiment has two scopes:

- Use *PlusMe* to detect warning signals of ASD;
- Correlate the behaviours with *PlusMe* with Modified Checklist for Autism in Toddlers (M-CHAT)<sup>1</sup>, a screening tool used to evaluate the possibility of autism.

The experiment plans a single session with a specific protocol to detect red flags for ASD (e.g., poor eye contact, low pointing, low level of imitation).

### 2.2. Participants

The experiment involves two sample of children:

- 1) TD group: a initial sample of 38 typically developing children (18 females and 20 males, aged between 12 and 31 months, mean age: 25 months) already evaluated during a previous project<sup>2</sup>. Based on the results by CRI (study described in sec. 4) the sample was reduced to 34 TD children (18 boys and 16 girls, aged between 20 and 31 months, mean age: 26 months). Children with neurodevelopmental disorders or disorders of an organic nature were excluded from the sample.
- 2) AR-ASD group (children At Risk of Autism Spectrum Disorder): children (paired to TD group by age and sex) undergoing diagnostic evaluation for suspected autism at

---

<sup>1</sup> <https://mchatscreen.com/>

<sup>2</sup> <https://www.istc.cnr.it/en/content/me>

Department of Human Neuroscience, Section of Child and Adolescent Neuropsychiatry, University of Rome *La Sapienza*. The inclusion criteria is a score above 5 (corresponding to a moderate/high level of risk for autism) on the Autism Diagnostic Observation Schedule-2 (ADOS-2)<sup>3</sup> Toddler Module. To date, we evaluated two children with AR-ASD: a male subject of 26 months and a female subject of 26 months.

The Ethics Committee of the National Research Council of Italy<sup>4</sup> approved the study with TD children (protocol No. “0039228/2019”), and the National Ethics Committee<sup>5</sup> approved the study with AR-ASD children (Protocol No. “0027279 Class: PRE BIO CE 01.00”, date 13/07/2022). The parents gave written informed consent before starting the experiment.

## 2.3. Procedure

### 2.3.1. Experimental setting

The TD group was recruited and tested at kindergartens of Rome. The AR-ASD children were recruited and tested at the Department of Human Neuroscience, Section of Child and Adolescent Neuropsychiatry, University of Rome *La Sapienza*.

Each child was tested individually in the presence of two people: the neurodevelopmental therapist, who plays with the child using the *PlusMe* toy, and an experimenter in charge of managing the control tablet; children didn't know them before. The experiment took place on a carpet, where there was the experimenter, the child, the *PlusMe* and several toys that the child knows (e.g., lego, doll, etc). During the experimental session, the therapist proposed different play activities based on the protocol.

The experiment had an overall duration of about 15-20 minutes. Before the experimental session, the parents (both mom and dad) and the child's teacher (for TD children) were requested to complete the M-CHAT, a standardised questionnaire consisting of 23 yes/no questions, used for early detection of possible warning signals in the communicative and relational competences (often critical symptoms of ASD).

### 2.3.2. Experimental protocol

The child and therapist sat down on the carpet. The *PlusMe* was turned off. At this moment, the experiment began. The therapist proposed seven different play activities, based on *PlusMe* operating modes:

1. **Novelty valuation phase (1-5 minutes):** The child is free to play with the toys present on the carpet, and the therapist plays with the child. If the child chooses *PlusMe*, the other toys can be removed and the protocol starts. In the case in which the child doesn't

---

<sup>3</sup>C. Lord, M. Rutter, P. C. Dilavore, S. Risi, K. Gotham, and S. L Bishop. 2012. (ADOS2) Autism Diagnostic Observation Schedule, Second Edition. Western Psychological Services, Los Angeles.

<sup>4</sup> Research Ethics and Integrity Committee, <https://www.cnr.it/en/ethics>

<sup>5</sup> [www.iss.it/en/comitato-etico](http://www.iss.it/en/comitato-etico)

choose *PlusMe*, the child and therapist play with different toys. After a couple of minutes, the experimenter activates *PlusMe* with rewarding patterns composed of different lights and sounds. If the child's attention is immediately captured, the therapist removes the other toys and starts to play with *PlusMe* (moving on to the next phase); otherwise (child's attention not immediately captured), the reward triggering is repeated up to three times. After three attempts, this phase is in any case finished and the toys on the carpet are removed. In this case, the therapist entices the child to play with *PlusMe*. In this phase the focus on novelty is assessed.

2. **PlusMe exploration phase (3 minutes):** In this activity, the child's ability to point to the correct part of *PlusMe* is assessed. The therapist asks the child to point to, for example, the ears, the eyes of the *PlusMe*. The therapist's questions can be: "Where is the nose?" "Where is the mouth?".
3. **Imitation phase (40 seconds):** The up-left paw of *PlusMe* is set in green and with the cow sound. The therapist touches the paw of *PlusMe* and entices the child to do the same (20 seconds). After the time, the therapist simultaneously touches the two up paws of *PlusMe* and entices the child to do the same (20 seconds). In this phase the ability to imitate the therapist's behaviours is assessed.
4. **Symbolic play phase (3 minutes):** The therapist cuddles *PlusMe* and entices the child to do the same (1 min and 30 sec). Then, the therapist invites the child to feed *PlusMe* (1 min and 30 sec). When the child correctly performs the behaviour, the experimenter activates a rewarding pattern (*PlusMe* emits a sound). The child's ability to execute symbolic behaviour is assessed.
5. **Turn-taking phase (2 minutes):** During this phase, a random paw emits a blinking red light; if it is touched, a rewarding sound is emitted (trumpet notes) and the colour turns green. After a couple of seconds, the game restarts with another random paw. The therapist invites the child to play chasing the blinking paw alternating the turn ("Now it's my turn, now it's your turn"). In this phase, the child's ability to play respecting the game turn is assessed.
6. **Tablet introduction phase (around 1 min and 30 sec):** The therapist points to the tablet in the hands of the experimenter and asks the child what it is. The experimenter gives the tablet to the therapist, who presses a button on the tablet to obtain a reward pattern on *PlusMe*. In this phase, the child is free to press the button to evaluate if understands the association tablet-*PlusMe*. At this stage, the child's ability to understand cause and effect is assessed.
7. **Turn-taking between children (2 minutes):** this is the same activity described above ("Turn-taking phase"), but this time two children play together, alternating the turn during the game.

## 2.4. Data collection

The experimental session was recorded with two cameras, and later analysed through a video-editing software. The clips were then rated to extract both duration (in seconds) and

frequencies of 8 behavioural indexes. In order to achieve the second scope (described in sec. 2.1), each index correspond to given items of M-CHAT:

1. *Pointing* (during *PlusMe exploration* phase): how many times the child points at something. This index corresponds to items 6 and 7 of the M-CHAT:

item 6: Does your child ever use his/her index finger to point, to ask for something?

item 7: Does your child ever use his/her index finger to point, to indicate interest in something?

2. *Explore* (during all phases): it evaluates if the child explores the device correctly, or if the child lies down and observes only a part of the device and does not have a complete view of *PlusMe*. This index corresponds to item 8 of the M-CHAT:

item 8: Can your child play properly with small toys (e.g. cars or blocks) without just mouthing, fiddling, or dropping them?

3. *Eye contact* (during all phases): how long time the child looks at the therapist, for more than 1 second. This index corresponds to item 10 of the M-CHAT:

item 10: Does your child look you in the eye for more than a second or two?

4. *Imitation* (during *imitation* phase): how many times the child correctly reproduces the therapist's behaviour on the toy. This index corresponds to item 13 of the M-CHAT:

item 13: Does your child imitate you?

5. *Smile* (all phases): how many times the child smiles at the therapist. This index corresponds to item 12 of the M-CHAT:

item 12: Does your child smile in response to your face or your smile?

6. *Symbolic play* (during symbolic play phase): how many times the child caresses and feeds the *PlusMe* in a correct way. This index corresponds to item 5 of the M-CHAT:

item 5: Does your child ever pretend, for example, to talk on the phone or take care of a doll or Yes No pretend other things?

7. *Name* (all phases): it evaluates the child's response to the name. This index corresponds to item 14 of the M-CHAT:

item 14: Does your child respond to his/her name when you call?

8. *Sequences* (during novelty valuation phase and tablet introduction phase): In the novelty valuation phase, when *PlusMe* is activated, it is assessed if the child looks at *PlusMe* and then at the therapist. In the tablet introduction phase, it is assessed if the child

directs the gaze at the tablet, the PlusMe and after the therapist. This index corresponds to the item 23 of the M-CHAT:

item 23: Does your child look at your face to check your reaction when faced with something unfamiliar?

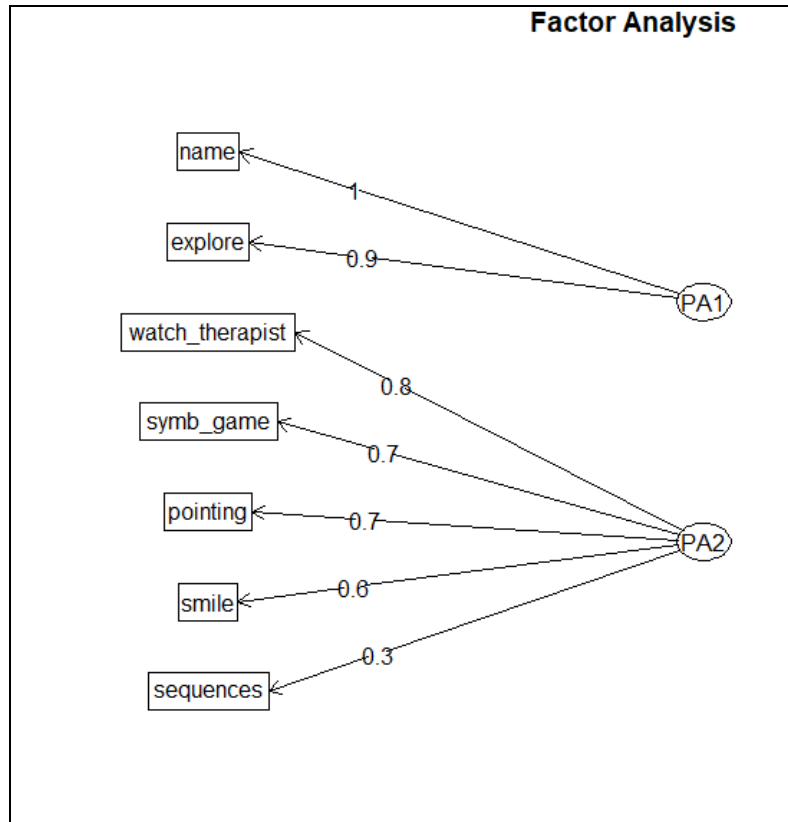
## 2.5. Preliminary results

The analysis was conducted on 21 children<sup>6</sup>. To reduce the number of indexes a factor analysis was conducted. A first examination of the Kaiser-Meyer Olkin (KMO) suggested that the sample, in the current form, wasn't factorable (KMO = 0.58); a subsequent analysis of the Measure of Sampling Adequacy (MSA) for each item suggested that the *imitation* index hasn't adequacy in the model (MSA = 0.38). Removing the *imitation* index, the KMO suggested that the sample was factorable (KMO = 0.61). The analysis yielded a two-factor solution (Figure 1). Using the factor loadings, two indexes are part of Factor 1 (PA1 in the figure), called *Environmental Interaction: explore* (0.94) and *name* (1.0); Five indexes are part of Factor 2 (PA2 in the figure), called *Social Interaction: pointing* (0.7), *sequences* (0.30), *smile* (0.60), *eye contact* (0.80), *symbolic play* (0.70). The *Environmental Interaction* factor refers to how the child relates and interacts with the outside world and responds to environmental stimuli. The *Social Interaction* factor refers to the way the child interacts with the other people, communicating and responding to relational intentions, needs, and meanings.

---

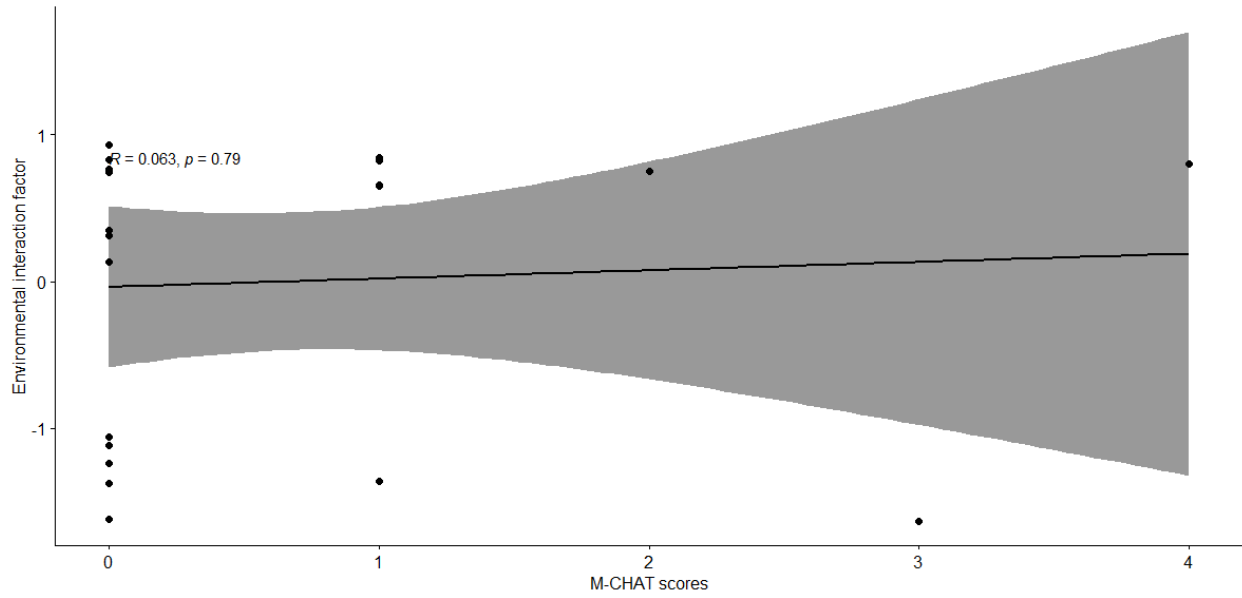
<sup>6</sup> The analysis is currently ongoing on the remaining participants.



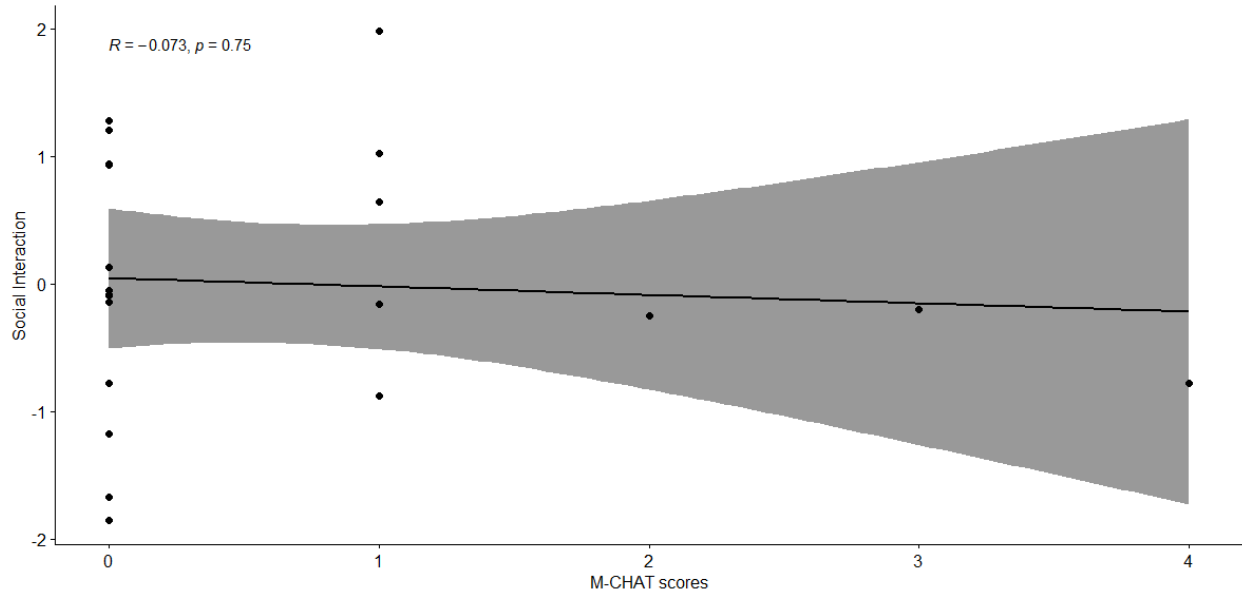


**Figure 1.** A diagram to schematize the number of factors and which indices make them up; for each arrow the factor loading is provided.

The next step will be correlating the *Social Interaction* and *Environmental Interaction* factors scores with the M-CHAT scores. A first preliminary plot suggests, as we expected, that the scores of TD children are very close together (Figure 2 and 3); we expect that the next data from AR-ASD children, to be opposed to TD children, should create two distinct groups, potentially useful for early detection of warning signals of autism.

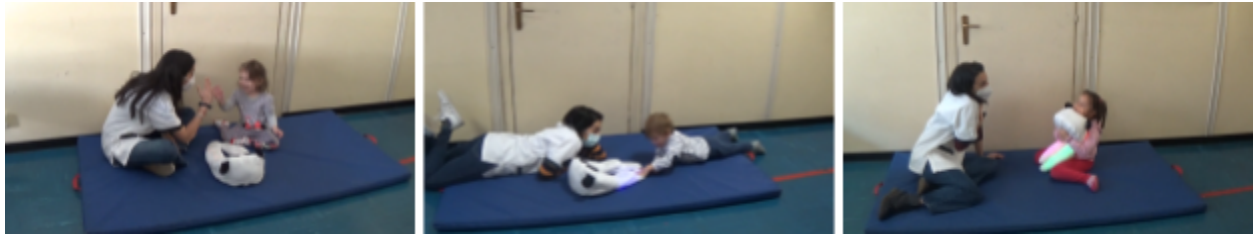


**Figure 2.** In the figure, each TD child (black dots, N=21) is plotted according to the M-CHAT and *Environmental Interaction* factor score.



**Figure 3.** In the figure, each TD child (black dots, N=21) is plotted according to the M-CHAT and *Social Interaction* factor score.

## 3. LA SAPIENZA: PlusMe as a therapeutic tool



### 3.1. Aim of the experiment

The *PlusMe* device is used to assess its potential therapeutic use during the early intervention in autism. The experiment plans four consecutive sessions and aims to validate the toy as a therapeutic tool; the objective is to see if repeated activities, based on *PlusMe*, promote an increase of key social behaviours critical in ASD<sup>7</sup>.

### 3.2. Participants

An initial sample of 26 children was recruited. Of these, three were excluded for a diagnosis differently of ASD, 13 children were excluded because they did not complete the four consecutive sessions, and one child was excluded because performed the activity for less than 75%. The final sample consisted of nine children, mean age 42 months, range 36-53 months. The acceptance criteria are based on a diagnosis of autism spectrum disorder based on DSM-5 criteria<sup>8</sup>. The participants were denoted as high-functioning subjects, presenting a moderate level of symptoms, according to the ADOS-2. The Ethics Committee of the National Research Council of Italy approved the study (protocol No. “0052419/2021”<sup>9</sup>), and the parents gave written informed consent before starting the experiment.

### 3.3. Procedure

The participants were recruited and tested at the Department of Human Neuroscience, Section of Child and Adolescent Neuropsychiatry, University of Rome *La Sapienza*. The experiment took place in an observation room where the distracting elements were removed (e.g., pens, other toys). Each child was tested individually for four consecutive sessions, one per week, in the presence of two people: the neurodevelopmental therapist, who played with the child using the *PlusMe* toy, and an experimenter in charge of managing the control tablet; children didn't know

<sup>7</sup> Here it is possible to see selected clips from a pilot experiment run on 3 children (age 36,42 and 50 months), diagnosed with high-functioning ASD <https://www.plusme-h2020.eu/video/>

<sup>8</sup> American Psychiatric Association. 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (5th ed.). Washington DC. 636–638 pages. <https://doi.org/10.1176/appi.books.9780890425596>

<sup>9</sup> This ethical clearance concerns the *PlusMe* European project

them before. During the experimental session, lasting around 10 minutes, the therapist proposed five different play activities (based on different *PlusMe* operating modes):

1. **exploratory activity:** each paw of *PlusMe* emits a different colour when it is touched. This activity is always the first;
2. **whack a mole activity:** a random paw emits a blinking red light; if it is touched, a rewarding sound is emitted (trumpet notes) and the colour turns to green. After a couple of seconds the game restarts with another random paw.
3. **caress activity:** if the child cuddles *PlusMe*, it emits a rewarding pattern (triggered by the experimenter through the control tablet);
4. **two hands activity:** if the upper paws are touched together, they light up in green and a brief sound is emitted (electronic ding).
5. **freedom activity:** the therapist asks the child his/her favourite *PlusMe* outcomes. The toy operating mode is then changed in real-time by the experimenter holding the tablet, on the basis of the child's preferences, expressed by both verbal and non-verbal communication (namely social request). In absence of a child's request, the therapist proposes an output to the experimenter. Notably, the therapist actively encourages the child's social engagement.

The activities with *PlusMe* help to promote different social behaviours such as attention, joint attention, emotional and imitation responses, motor coordination, and explicit social requests.

### 3.4. Data collection

The experimental sessions were recorded with two cameras, and later analysed through a video-editing software. The exploratory activity was not taken into consideration for the data analysis. The clips were then rated to extract both duration (in seconds) and frequencies of different behavioural indexes:

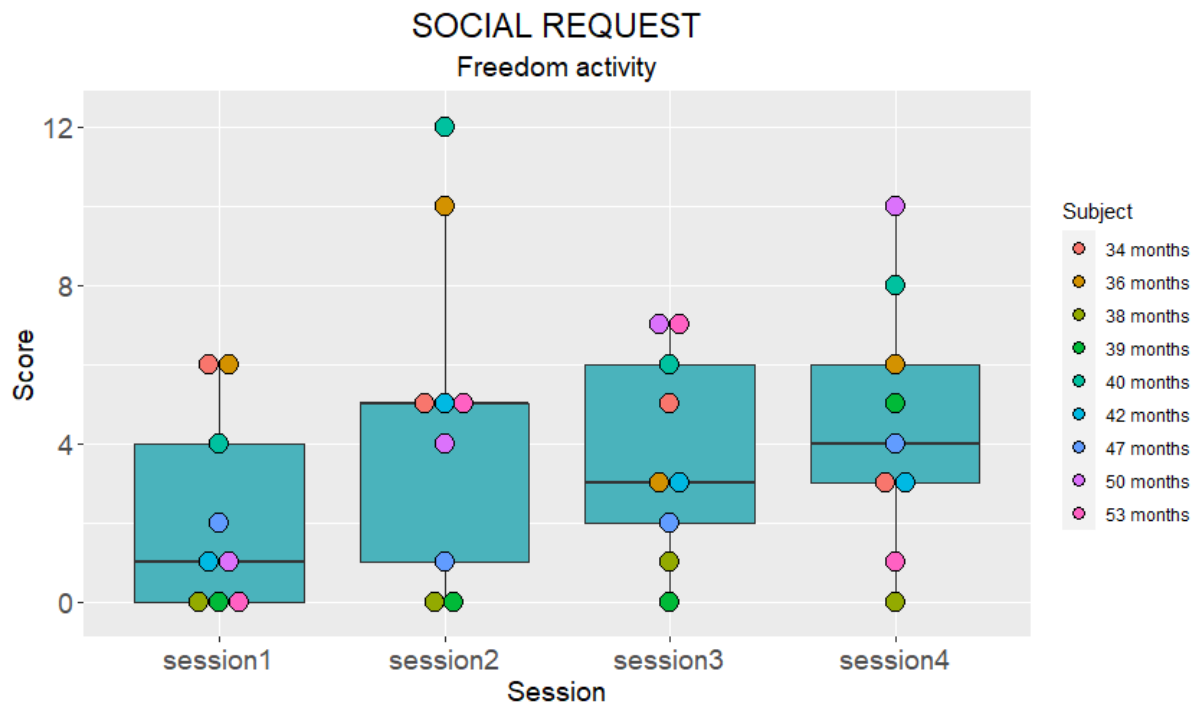
- *imitation:* how many times the child correctly reproduces the therapist's behavior on the toy during caress and two hands activities;
- *watch therapist:* how long time the child looks at the therapist during all activities (all activities);
- *smile:* how many times the child smiles at the therapist (all activities);
- *social request:* how many times the child asks the therapist or the experimenter — verbally or not — to change the rewarding pattern of *PlusMe* during freedom activity;
- *watch PlusMe:* how long time the child looks at *PlusMe* (all activities);
- *sequences:* how many times the child looks first at the *PlusMe* and after the therapist (all activities).

Such indexes have been selected as they provide a general idea of the social interaction between child and therapist. In this deliverable only the following 3 indexes have been analysed

<sup>10</sup>: *watch therapist, social request and imitation*. To see if the repeated therapy sessions based on *PlusMe* improves some social behaviour, we run a Wilcoxon signed-rank test, comparing the results from first and fourth sessions.

### 3.5. Results and discussion

The Wilcoxon signed-rank test shows a statistically significant difference in the *social request* index in session 1 and session 4 ( $p = 0.05$ ) (figure 4). This indicates that the child, during the four session, increased the social interaction with the therapist or experimenter, and asked them to change the output of *PlusMe*.

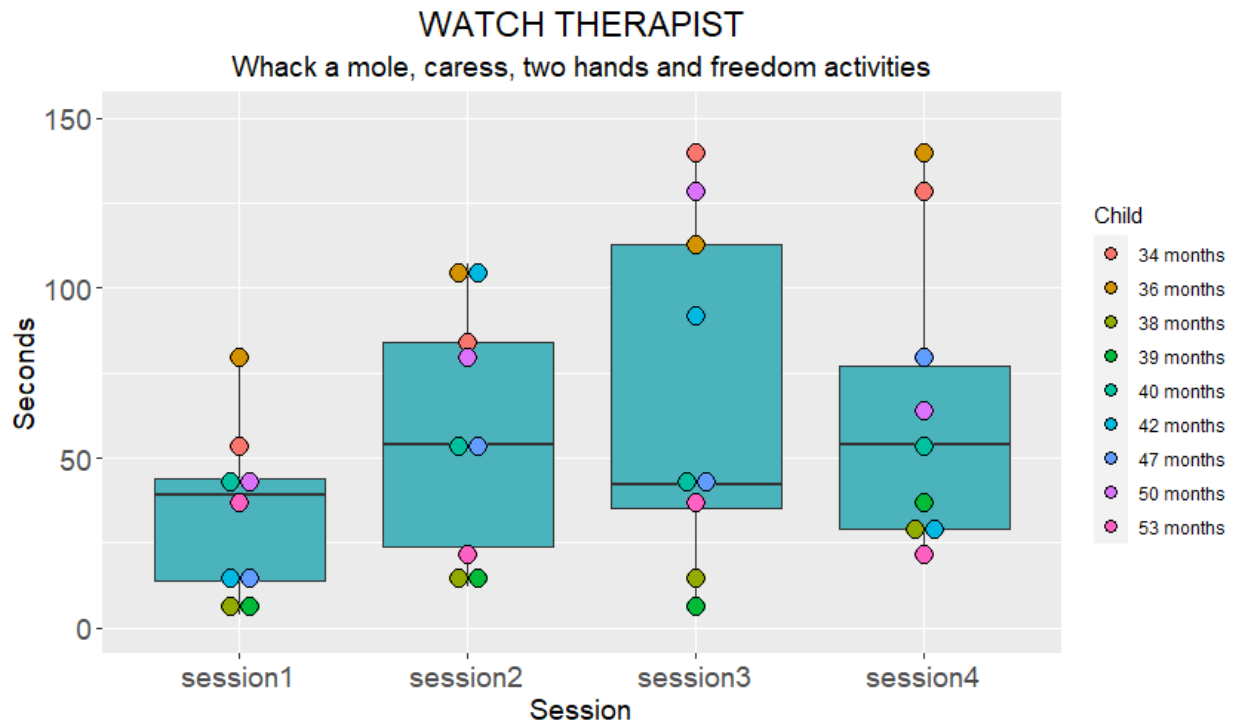


**Figure 4.** Box plot referred to the number of *social requests* made by children throughout the four sessions, during the *freedom* activity. Each dot represents a child. The difference between 1 and 4 sessions is statistically significant ( $p=0.05$ ).

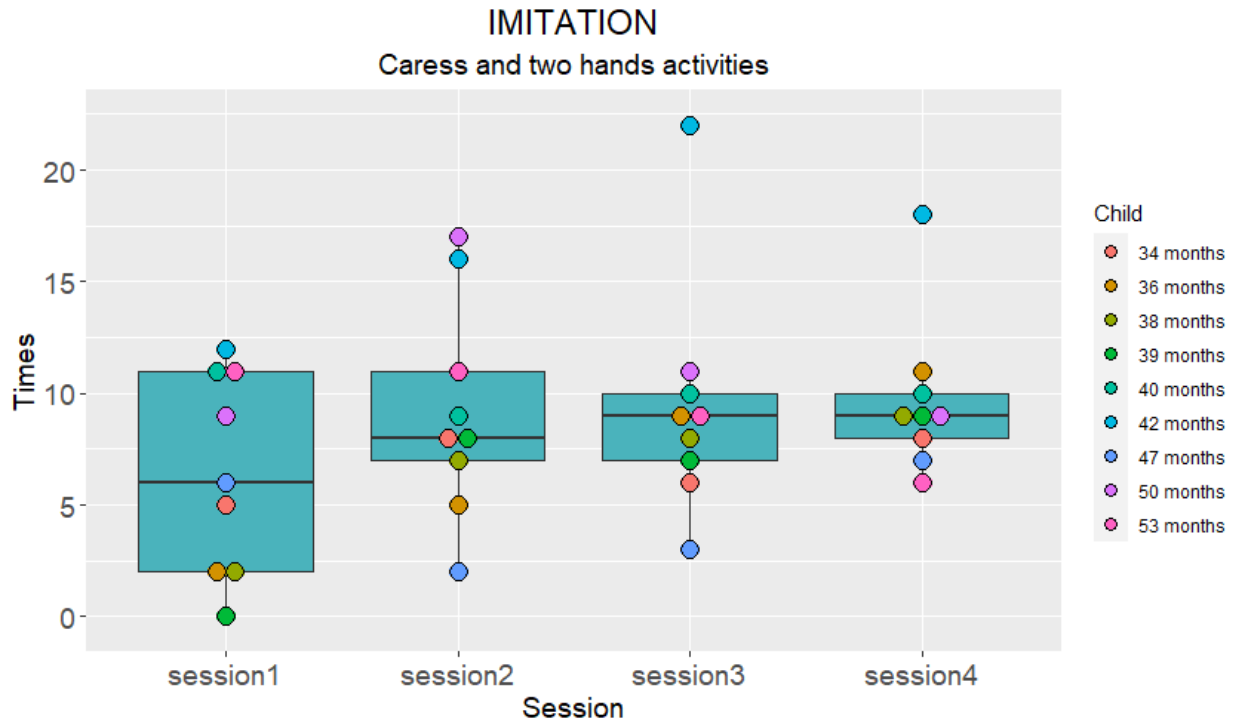
The Wilcoxon signed-rank test conducted on the *watch therapist* index is statistically significant ( $p = 0.01$ ). The box plot in figure 5 shows how children spend more and more time looking at the therapist in all activities. This is an important observation given the poor eye contact behaviour evident in the participants, a key clinical symptom of ASD. The Wilcoxon signed-rank test conducted on the *imitation* index was statistically significant ( $p = 0.04$ ). The box plot in figure 6 indicates how the participants' ability to correctly imitate the therapist's behaviour, during the *caress* and *two hands* activities, improves throughout the sessions. Qualitatively the children

<sup>10</sup> The analysis of the remaining indexes is currently ongoing

seem to pay more and more attention to the therapist's gestures to trigger the *PlusMe* rewarding pattern.



**Figure 5.** Box plot referred to *watch therapist* index during all activities, throughout the four sessions. The difference between 1 and 4 sessions is statistically significant ( $p=0.01$ ).



**Figure 6.** Box plot of the *imitation* index during the caress and two hands activities, throughout the four sessions. The difference between 1 and 4 sessions is statistically significant ( $p=0.04$ ).

## 4. CRI: PlusMe and the early diagnosis of ASD

### 4.1. Aim of the experiment

The purpose of this study was to determine the age range at which *PlusMe* can be used as a diagnostic tool for ASD. More precisely, the experiment sought to determine from what age to what age TD children demonstrate key behaviours during play with *PlusMe*, whose absence would suggest a suspicion of ASD (see Section 2 for more details on those behaviours).

### 4.2. Participants

For this study, children were recruited in 3 crèches of Calais, a city in the North of France. 47 children participated in the study, 27 boys and 20 girls, aged between 12 and 45 months.

The exclusion criteria were: being screened or being in the process of being screened for ASD (N=1) and being at 'medium' or 'high' risk of ASD as a result of responses to the French version of the M-CHAT-R (N=3). In addition, 9 children refused to participate in the experiment because of shyness (N=5) and fear (N=2) or because their attention was diverted by the surrounding noise (N=2).

The final sample consisted of 28 children, 12 boys and 16 girls, aged between 12 and 45 months (N=28), with a mean age of 26.18 months (median = 27.5; standard deviation = 8.68).

## 4.3. Procedure

### 4.3.1. Experimental setting

The experiment took place in an observation room (see Figure 7). Each child was tested individually for 15-20 minutes, with a specific protocol, in the presence of two experimenters: one, who played with the child using the *PlusMe* toy, and the other in charge of managing the control tablet. Children had not previously met the experimenters. The experiment took place at a table. At the table, there was the experimenter, the child, the *PlusMe* and several toys that the child knew before the experiment (e.g., lego, doll, etc ...). During the experimental session, the experimenter proposed different play activities based on the protocol (see next section).

The experiment had an overall duration of about 15-20 minutes. After the experiment, early childhood educators of the crèche were requested, for each child, to complete the French Version of the Modified Checklist for Autism in Toddlers (M-CHAT-R).



Figure 7. Experimental setting.

### 4.3.2. Experimental protocol

The experimental protocol used was the same as that described in Section 2.3.2. However, in this study, the last activity (Turn-taking between children) was not carried out because the experimental conditions in the crèches did not allow it.



## 4.4. Data collection

The experimental sessions were recorded with a 360° camera, and later analysed using video-editing software. Two experimenters independently coded the videos to determine "*PlusMe* score" and "*PlusMe* adaptability" for each child.

### ***PlusMe* score (scored between 0 and 6)**

For each child, each of the six activities was coded as success (1) or failure (0). Specifically, we adopted the following criteria:

- **The "Exploring novelty" activity** was coded as 1 when the child, at a minimum, observed the *PlusMe* with attention and touched it.
- **The "Pointing at the *PlusMe*" activity** was coded as 1 when the child pointed in the direction of the *PlusMe*, at least 3 times when he/she was asked to do so.
- **The "Imitation" activity** was coded as 1 when the child imitated the action performed by the experimenter with the *PlusMe* (press on both front paws simultaneously) - a verbal guidance could be given to the child, such as "Do the same as I did!"
- **The "Symbolic play" activity** was coded as 1 when the child fed *PlusMe* with plastic utensils that were placed on the table (plates, cutlery, plastic food), but also when the child fed *PlusMe* with toys that were not specifically offered for this activity.
- **The "Turn taking" activity** was coded as 1 when the child respected the turn-taking between him/her and the experimenter to press the lighting paw of the *PlusMe* - a verbal cue could be given to the child, such as "It's my turn! It's your turn!". The activity was also coded 1 when the child intentionally transgressed the turn-taking pattern; it means that the child knew it was not his/her turn to play, but the experimenter's turn, and he/she pressed the *PlusMe* paw anyway for fun.
- **The "Joint attention/introduction to the tablet" activity** was coded as 1 when the child alternately watched the experimenter controlling the tablet, the tablet and the *PlusMe*.

### ***PlusMe* adaptability**

In addition to the "*PlusMe* score" variable, a second variable was used to reflect, for each child, the experimenters' feelings about the degree of appropriateness of the *PlusMe* and the activities proposed. This variable, "*PlusMe* adaptability", included the following four levels:

- **Unadapted toy:** Children who, during *PlusMe* activities, did not seem to understand the activities, even when instructions were given and guidance provided. We classified in this category children who easily diverted their attention to other stimuli and showed signs of willingness to return to the collective living environment.
- **Adapted toy (-):** Children who during *PlusMe* activities seemed to understand the majority of the activities to be performed, but for whom verbal instructions and guidance were often required.
- **Adapted toy (+):** Children who, during *PlusMe* activities, understood all the activities to be performed and for whom very few verbal instructions were given. Children in this

category showed the most interest and enjoyment in interacting with the *PlusMe* and the experimenter.

- **Obsolete toy:** Children who, during *PlusMe* activities, understood all the activities, without any instructions or guidance. Children in this category appeared to have little interest in *PlusMe* and showed signs of boredom.

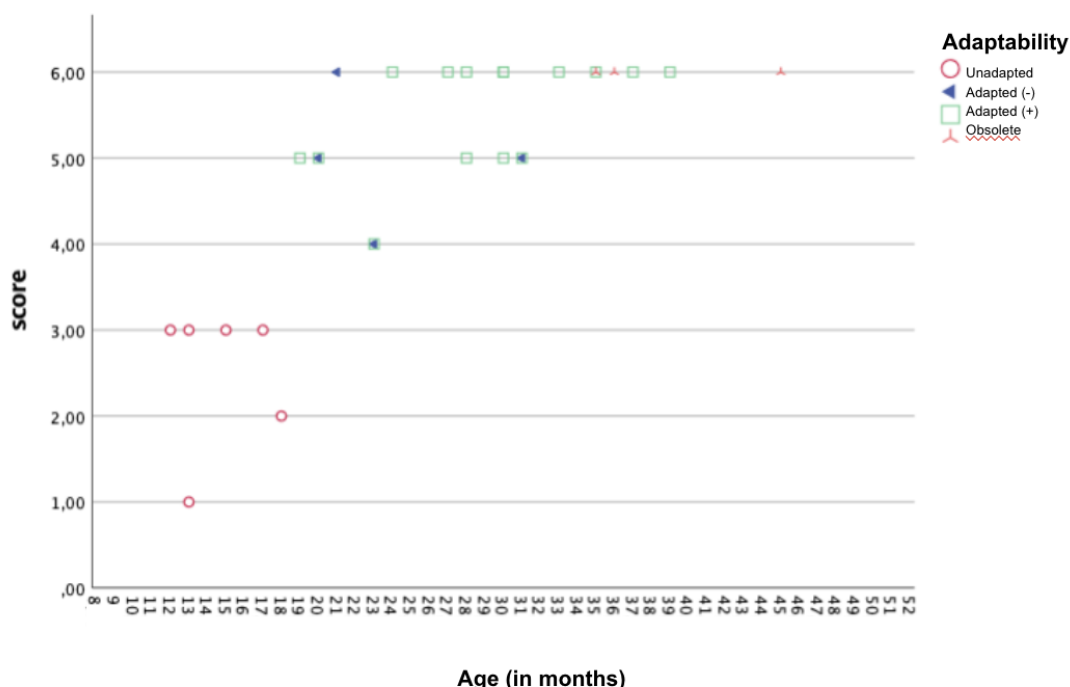
## 4.5. Results and discussion

In order to determine from what age to what age TD children playing with *PlusMe* demonstrate key behaviours whose absence would suggest a suspicion of ASD, we first observed the distribution of the variable “*PlusMe* score” according to age (in months) and “*PlusMe* adaptability” categories (see Figure 8).

In Figure 8, we observe that :

- For the "**Unadapted toy**" category (n=6), the age range was 12 to 18 months. None of the children had a score higher than 3, four children had a score equal to 3, one had a score equal to 2 and one had a score equal to 1.
- For the "**Adapted toy (-)**" category (n=4), the age range was 20 to 31 months. One child had a score of 6, two had a score of 5 and one had a score of 4. This implies that 75% of the children had a score of 5 or higher.
- For the "**Adapted toy (+)**" category (n=15), the age range was 19 to 39 months. Nine children had a score of 6, five had a score 5 and one scored 4, implying that 93% of the children scored 5 or higher.
- For the "**Obsolete toy**" category (n=3), the age range was 35 to 45 months. 100% of the children had a score equal to 6.

These observations are summarised in Table 1.



**Figure 8.** Distribution of the variable “*PlusMe* score” according to age (in month) and “*PlusMe* adaptability” categories.

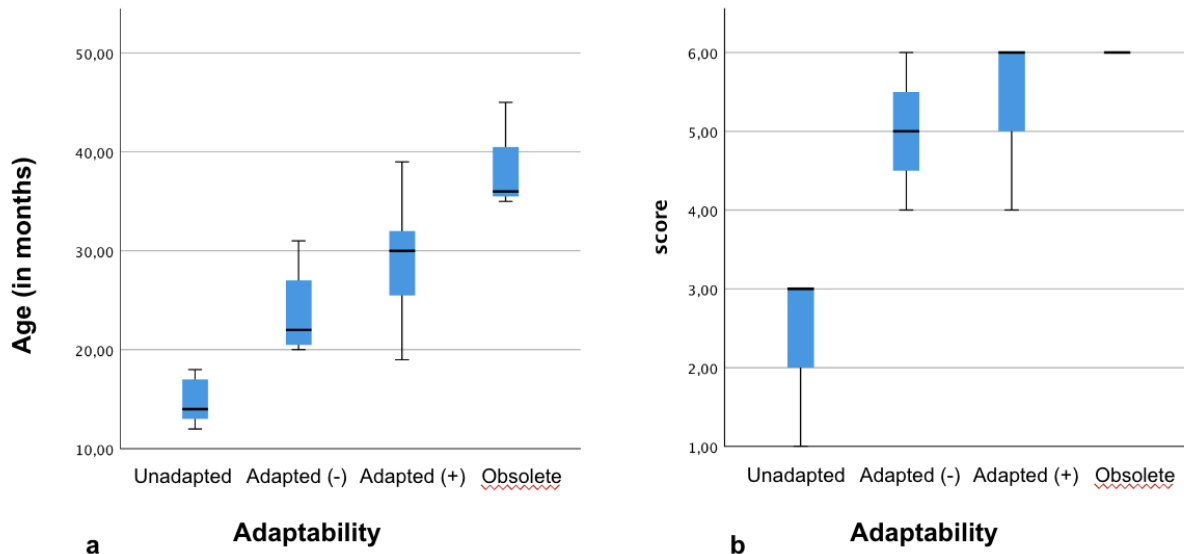
Adaptability	N	Min age (mo)	Max age (mo)	Median age (mo)	Median <i>PlusMe</i> score
Unadapted	6	12	18	14.0	3.00
Adapted (-)	4	20	31	22.0	5.00
Adapted (+)	15	19	39	30	6
Obsolete	3	35	45	36	6

**Table 1.** For each adaptability category, number of children, minimum age, maximum age, median age (in months) and median *PlusMe* score.

We performed a Kruskal-Wallis test with factors: age and adaptability, showing a statistically significant difference ( $Kw=17.967$ ;  $p<.001$ ) in the median ages between the different adaptability categories (see Figure 9a). Dunn's post hoc test with Bonferroni adjustment showed a statistically significant difference in the median age between the "Unadapted" and "Adapted (+)" categories after Bonferroni adjustment ( $Kw=13.7$ ;  $p=.003$ ), and between the "Unadapted" and "Obsolete" categories ( $Kw=18.5$ ;  $p=.004$ ).

We performed a Kruskal-Wallis test with factors: *PlusMe* score and adaptability, showing a statistically significant difference ( $Kw=17.649$ ;  $p<.001$ ) in the median scores between the different adaptability categories (see Figure 9b). Dunn's post hoc test with Bonferroni adjustment showed a statistically significant difference in the median “*PlusMe* score” between

the "Unadapted" and "Adapted (-)" categories after Bonferroni adjustment ( $Kw=14.2$ ;  $p=.001$ ) and between the "Unadapted" and "Obsolete" categories ( $Kw= 18.5$ ;  $p=.004$ ).



**Figure 9.** (a) Age (in months) according to the Adaptability category ; (b) *PlusMe* Score according to the Adaptability category.

Based on this first study, we conclude that the appropriate age range for the *PlusMe* to be used as an early screening tool for ASD would be between **18 and 32 months of age**. However, in order to draw more definite conclusions, it would be necessary to conduct a larger-scale study, including the same number of children in each “adaptability” category.

## 5. Future Developments

The next deliverable plans the empirical validation of the whole IM-TWIN system, designed by:

- sensorised t-shirt, to detect the engagement state of the child;
- computer vision, to analyse the gaze from child to therapist.

This empirical validation will be given in the next deliverable D4.2 Empirical validation: IM-TWIN (due al 31 Oct 2023), in which the results of the empirical validation of the whole IM-TWIN system as a toy for therapy, and for the early diagnosis will be given.

### 5.1. New experimentation and testing

This section contains a description of the new experiments conducted and to be conducted to test the new prototype of *PlusMe* and App, the computer vision with camera glasses, the new *Octopus X-8* device in a real-life scenario, and the experiment with sensorised t-shirt.

### 5.1.1. PlusMe and App

As described in D3.6 PlusMe production 2<sup>11</sup> and in the PlusMe project deliverable D1.3 PlusMe product demonstrator<sup>12</sup> an updated version of *PlusMe* toy and App is available. This new version of *PlusMe* and the App has been tested on three children with neurodevelopmental disorders at the Department of Human Neurosciences, Section of Child and Adolescent Neuropsychiatry (University of Rome *La Sapienza*); plans are in place to recruit 10 more children.

The first results showed that the possibility to select more colours and the paw vibration engaged the children much more, who made more social requests (i.e. to change the output of the *PlusMe*). With the new update and the possibility of selecting 10 different colours, it is possible that the child will request its favourite colour and that it can be selected on the *PlusMe*.

Furthermore, the results of the use of the new version of PlusMe App have shown that it is easier to use and understand, and very user-friendly; a questionnaire to investigate this aspect is underway and will be delivered in the deliverable D5.12 Country-based exploitation questionnaire and stakeholder interviews.

In addition, the new *PlusMe* device and App will be used in Paris by CRI for a new experiment (see next section).

#### 5.1.1.2. Experiments with new *PlusMe* (CRI)

The aim of these experiments will be to develop a scoring system for the behaviours (such as pointing, imitation, joint attention) exhibited by a child during *PlusMe* activities, allowing the therapist to assess the developmental skills that need to be encouraged in a given child. To this end, we will conduct a two-stage research study: the establishment of a scoring system for behaviours observed during *PlusMe* activities with typically developing children (TD) and the evaluation of the use of this scoring system with children with ASD.

Study 1. This study was conducted with 28 children with TD, aged 12 to 45 months (12 boys, mean age = 26.18 months). Two experimenters went to the nursery to carry out individual sessions around 6 activities with *PlusMe*. An analysis of the video recordings will allow the creation of a scoring system specifying for each activity: (i) the stereotyped behaviours of the child that specifically respond to the instruction (e.g., using a spoon for the instruction "Feed *PlusMe*") and (ii) the non-stereotyped elaboration behaviours produced spontaneously by the child, which go beyond the instruction (e.g., imagining having an ice cream for the same instruction).

Study 2. This study will be conducted with ten children with ASD and ten children with TD, aged 3 to 6 years. Two experimenters will carry out individual sessions around the 6 activities with *PlusMe*. During the test, the behaviours presented by the child will be scored using the scoring system presented above. This will make it possible to establish a score for each child according to two dimensions: stereotyped behaviour and spontaneous elaboration behaviour. We

<sup>11</sup> [https://im-twin.eu/wp-content/uploads/2022/11/DELIVERABLE\\_D3.6\\_PlusMe\\_production\\_2.pdf](https://im-twin.eu/wp-content/uploads/2022/11/DELIVERABLE_D3.6_PlusMe_production_2.pdf)

<sup>12</sup> [www.plusme-h2020.eu/wp-content/uploads/2021/12/D1.3\\_plusme\\_product\\_demonstrator.pdf](http://www.plusme-h2020.eu/wp-content/uploads/2021/12/D1.3_plusme_product_demonstrator.pdf)

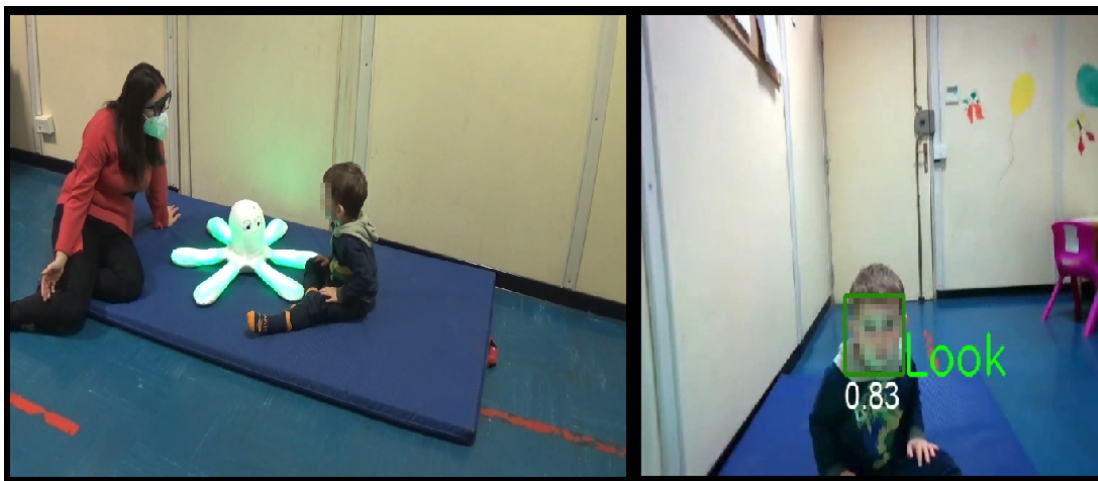
hypothesise that children with ASD will score as well as children with TD on stereotyped behaviours (previously acquired) but will score significantly lower than children with TD on spontaneous elaborative behaviours.

If our hypothesis is validated, we would like to make *PlusMe* available to healthcare professionals working with children with ASD under the age of 6, so that this 'intelligent' companion can support them in the care of the children and more specifically in the identification of non-stereotyped developmental skills that need to be encouraged in the child.

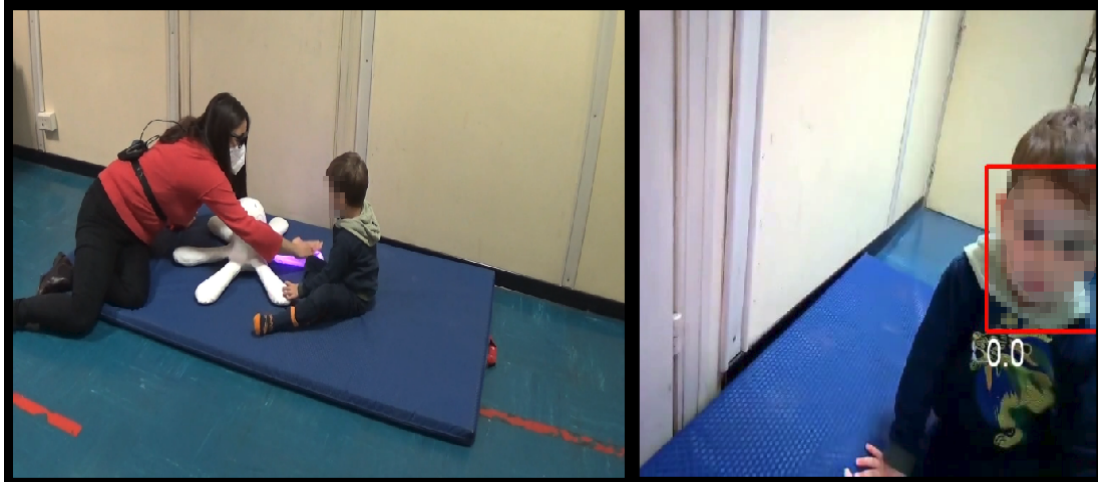
### 5.1.2. Camera glasses

As described in D2.1 Processing of physiological signals, visual information, and PlusMe interaction: first version (section 3)<sup>13</sup>, the CNR-ISTC developed a system for the automatic detection of eye contact between child and therapist. This behavioural index is extremely important, both in therapy and during the assessment of diagnosis; the lack of eye contact in children indicates generally an atypical social behaviour, and it is one of the most evident signals of ASD. The camera glasses have currently been used on 6 children with neurodevelopmental disorders.

Our aim is to use the camera glasses in all future experiments with TWCs to analyse the onset of joint attention, i.e. whether, following a reward (i.e. switching on the device because the interaction has taken place correctly), the child looks at the therapist. Figure 10 shows an example of this study (in this pilot the camera glasses and the *Octopus X-8*, described below, were used). In addition, the video recording through camera glasses will be used to collect dataset for facial expression recognition (for information see D3.3 PlusMe AI-augmented behaviour, and IM-TWIN 1).



<sup>13</sup>[https://im-twin.eu/wp-content/uploads/2023/02/DELIVERABLE\\_D2.1\\_Processing\\_of\\_physiological\\_signals\\_PlusMe\\_interaction.pdf](https://im-twin.eu/wp-content/uploads/2023/02/DELIVERABLE_D2.1_Processing_of_physiological_signals_PlusMe_interaction.pdf)



**Figure 10.** An example of output by camera glasses. (Up) When the child looks at the therapist, the output of artificial intelligence indicates this event with a green box and the probability that the child is actually looking into the eyes is shown; (Down) When the child doesn't look at the therapist, the output of artificial intelligence indicates this event with a red box and the probability that the child is actually looking into the eyes is shown.

### 5.1.3. Octopus X-8 device

As described in the PlusMe project deliverable D1.3 PlusMe product demonstrator<sup>12</sup> the CNR-ISTC researchers developed an additional interactive toy based on a modified version of the same hardware and software of *PlusMe*, called *Octopus X-8*. The device is shaped like an octopus, with six tentacles, and can detect a human touch on them by capacitive sensors, and emit coloured lights and sounds in response to that input, by embedded LEDs and speakers. One feature that makes this smart toy unique is that the toy can autonomously detect the identity of two Users. It can "understand" if its tentacles are touched by UserA or UserB, and then responds differently according to the User. This feature allows incrementing turn-taking games. The *Octopus X-8* is currently being tested on five children with neurodevelopment disorders; some results were presented during the workshop held 4 February, at the auditorium of the Department of Human Neurosciences, Section of Child and Adolescent Neuropsychiatry (University of Rome *La Sapienza*), and described in the deliverable D6.8 Workshop for therapists and rehabilitation centers <sup>14</sup>.

### 5.1.4. Experiment with sensorised t-shirt

CRI planned to start (around April) the experiment with the sensorised t-shirt in order to furnish UU suitable physiological data for TD children, necessary to extract three affective states (plus one extra state used as a baseline):

- **Positive engagement:** defined as high arousal and positive valence;
- **Negative engagement:** defined as high arousal and negative valence

<sup>14</sup> [https://im-twin.eu/wp-content/uploads/2023/02/DELIVERABLE\\_D6.8\\_workshop\\_for\\_therapists\\_1.pdf](https://im-twin.eu/wp-content/uploads/2023/02/DELIVERABLE_D6.8_workshop_for_therapists_1.pdf)

- **Low engagement:** defined as low arousal and both positive and negative valence;
- **Neutral engagement:** defined as middle arousal and middle valence.

The experiment will be conducted on 24 TD children aged 30-60 months. To verify the correct wearing of the t-shirt, the Signal Quality Indicator (SQI) developed by UU will be used (see D2.1 Processing of physiological signals, visual information, and PlusMe interaction: first version (section 2)<sup>13</sup>).