This is the accepted version of a paper presented at 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 29 August-2 September 2022, Napoli, Italy.

Citation for the original published paper:

Calvo-Barajas, N., Castellano, G. (2022)
Understanding Children’s Trust Development through Repeated Interactions with a Virtual Social Robot
In: 2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) (pp. 1451-1458). Institute of Electrical and Electronics Engineers (IEEE)
https://doi.org/10.1109/RO-MAN53752.2022.9900537

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-481074
Understanding Children’s Trust Development through Repeated Interactions with a Virtual Social Robot

Natalia Calvo-Barajas\textsuperscript{1} and Ginevra Castellano\textsuperscript{1}

Abstract—Studies in Child-Robot Interaction have shown that children form first impressions of a robot’s trustworthiness that might influence how they interact with social robots in long-term interactions. However, how children’s trust in robots evolves and how it relates to relationship formation is not well understood. This study investigates the effects of repeated encounters with a virtual social robot on children’s social and competency trust in social robots and their relationship formation. We developed an online storytelling game with the Furhat robot, where 25 children (9-12 years old) played with the robot over two sessions with seven days of zero exposure in between. Results show that children’s competency trust improved with time. We also found empirical evidence that children felt closer to the robot in the second encounter. This work enriches the scientific understanding of children’s trust development in social robots over extended periods of time in child-robot collaborative interactions.

I. INTRODUCTION

In settings such as children’s education or healthcare, socially assistive robots are becoming more present in digital and physical forms. Depending on the objective of these applications, many robotics systems are required to interact with children over extended periods. Hence, the understanding of children’s perceptions of social robots over time is gaining more attention in the Child-Robot Interaction (cHRI) community [1]–[3].

Previous studies have shown that children form first impressions of a robot’s trustworthiness in a few seconds [4]. Furthermore, children’s perceptions of trust have been assessed after interacting with a social robot in short-term interactions [5], [6]. However, these perceptions of trust in robots could be highly influenced by the novelty effect [7], and it is unclear whether children’s perceptions of social robots generalized across repeated encounters [8].

In addition, the existing literature on children’s trust in robots sets a distinction between two dimensions: social trust, which is related to the evaluation of a robot’s moral characteristics, and competency trust, which refers to a robot’s performance [8], [9]. In this line, an important question remains unanswered concerning long-term interactions with robots: How does children’s social and competency trust in social robots change over time?

With this work, we aim to address this question by providing empirical evidence of the effect of repeated interactions with a social robot on children’s trust development. We designed an online storytelling game with the virtual Furhat robot\textsuperscript{1}. Children were asked to play the game in two different sessions with seven days of zero exposure in between. We measured children’s trust beliefs, children’s perceived social and competency trust in the robot, and likability by using subjective measures; and relationship formation through objective measures.

Our work makes two main contributions. First, we found that children’s trust beliefs in friends and general robots follow the multifaceted nature of trust and have temporal consistency. This result is relevant to the cHRI community as it sets a baseline when evaluating trust in robots. Second, we provided empirical evidence that competency trust and relationship formation evolve with time. This research provides new insights and raises questions about the trust dynamics toward long-term cHRI.

II. RELATED WORK

A. Children’s Trust Development

Trust is crucial for successful interpersonal relationships between children and other people. It is formed since infancy and affects personal, social, and intellectual functioning through life [10], [11]. As children’s trust is highly personalized, it is subject to fluctuations based on previous experiences. In other words, trust is a dynamic process related to an individual’s beliefs and expectations about others [12].

Studies in psychology have shown that trust is a multidimensional construct complex to contextualize and measure. Literature suggested that two main trust dimensions emerge.

\textsuperscript{1}Virtual robot used in this study: \url{https://furhatrobotics.com/}
The first one relates to moral and interpersonal trust (e.g., benevolence, honesty), whereas the second one pertains to cognitive trust (e.g., performance, reliability) [13], [14].

Due to this multidimensional nature of trust, multiple factors impact children’s conceptualization of trust. For instance, children’s age [11], past relationships with others and personal attachment [15], and individual’s cognitive skills [16]. In middle childhood, children’s friendships play a key role in their development of trust when building relationships. Children assess several pieces of information at this age as their understanding of trust matures and set trust as a fundamental basis of friendships [11], [17]. There is evidence that children in second-and-third grades have temporal consistency of trustworthy behaviors across time and situations [18], which is crucial when understanding trust dynamics over time.

B. Trust in cHRI

The multidimensional nature of trust can be extended to Human-Robot Interaction (HRI) [19]. Across cHRI, the two main dimensions are also defined: (1) Social trust as the “belief that the robot will keep its word or promises”, and Competency trust as “perceived competency and reliability of the robot” [8]. Although trust requires time to establish, previous research has shown that children judge trust in robots in first encounters [4]. Therefore, it is important to understand whether trust increases or decreases over multiple repeated interactions.

While there is evidence that the length of time (e.g., short-and-long-term encounters) a user had been interacting with the robot has a significant influence on HRI trust [20], its investigation in cHRI is scarce. A recent meta-analysis found that the interaction length played a critical role on cHRI trust. Authors found that shorter interactions may result in higher competency trust. However, they could not assess how the trust dynamics changes over time since, at the time of their research, only two studies had investigated children’s trust in multiple interactions over multiple time points [8].

When children meet social robots for the first time, they tend to build new relationships with these unknown artificial agents [21]. For the robot to fulfill its goal as a social companion, interactions should rely on building trusted relationships between children and robots that are essential to long-term robot acceptance. Disclosing personal information is a behavior related to closeness and a predictor of children’s trust in other people [22], [23], which can also indicate how the relationship with robots develops [24]. In long-term interactions, a user study showed that children between 5 and 6 years old were more likely to disclose personal information to the robot in the post-test than in the pre-test [25]. Contrary, another study found that children between 8 and 10 years old felt closer to the robot after two months of interaction. However, children did not necessarily disclose more information after several encounters [3]. These inconsistent findings of the impact of time on children’s trust in and relationship with social robots appear to depend upon the stage of children’s development and the nature of the scenario.

C. The Novelty Effect in cHRI

One of the reasons why longitudinal studies are gathering more attention in cHRI relates to the notion of overcoming the novelty effect that is present in single encounters with social robots [7], [26]. Early work showed that children interacted more often with the humanoid robot “Robovie” in the first week of the interaction than in the second week, suggesting that children lost interest in the robot after building a model of the robot’s role and behaviors [27]. Furthermore, as initial expectations are formed, the novelty effect is shaped by the robot’s anthropomorphism and the familiarization with it [26].

Repeated encounters over multiple time points are associated with the idea of overcoming the novelty effect [3], [7]. However, understanding how children’s behavior changes as the novelty effect wear off and the effects on trust development remains unclear. The work presented in this paper is one of the few exploratory user studies investigating the multidimensional nature of children’s trust in robots over repeated interactions.

III. Research Questions

As discussed in Section II-B, many factors influence children’s judgment of trustworthiness in robots. Previous studies have shown that measuring trust in HRI is subject to the context and might change depending on the individual and their propensity to trust [28]. By studying children’s relationships with their friends, one can draw a starting point to identify how children conceptualize trust and their tendency to trust [11]. Therefore, we first aimed to understand how children trust their friends and general robots. We hence pose the following Research Question (RQ):

RQ1: How do children conceptualize the multifaceted nature of trust? And, Does this concept change over time?

To answer these questions, we assessed children’s trust beliefs towards their friends and general robots at the beginning and end of each session. The assessment yields a baseline to understand how children define and judge the multifaceted nature of trust, and the repeated measures give insights if these beliefs change over time.

Following up on our previous research [6], this study aimed to investigate how repeated interactions of a collaborative storytelling game could influence children’s perceptions of a social robot. We propose the following RQ:

RQ2: How do repeated encounters with a virtual social robot influence children’s perceptions of social trust, competency trust, and likability of the robot?

To answer this question, we used subjective measures to assess children’s perceptions of social trust, competency trust, and likability after interacting with the social robot.

Previous research has discovered that trust is a precursor to self-disclosure and closeness [22], and is a component of intimacy in children’s relationships with social robots. However, only a few works have focused on longitudinal
studies [2], [3]. We attempt to further our understanding of how repeated interactions with social robots might affect children’s development of relationships and rapport with robots. We aim to answer the following RQ:

RQ3: How do repeated encounters influence children’s relationship development with a virtual social robot?

Following findings in social psychology, we expected repeated encounters to increase intimacy and closeness in relationships with the robot.

IV. METHODOLOGY

We designed a user experiment with Interactive Session (sessions S1 and S2) as a within-subject factor. Participants interacted with a virtual storyteller social robot in the two sessions, with seven days of zero exposure.

A. Scenario

We built a simplified version of the storytelling game developed in [6]. Children created a story with the robot during the interactive storytelling game. To help children elaborate on their story plot, we designed a graphical interface using the PlayCanvas Web Game Engine. The game included three topics (i.e., a rainforest, a beach, and a farm) with two different scenes each. Each scenario had nine characters and eight objects that children could move through the various scenes using the mouse. Figure 2 illustrates an example of the graphical interface.

B. Virtual Robot

We used the virtual version of the Furhat robot, a human-like social robot with a virtual face back-projected on an embodied head with three degrees of freedom [29]. To resemble the child-like appearance of the robot, we adopted the default anime texture (Figure 3). We employed text-to-speech to reproduce the robot’s speech by using the Amazon Polly female Spanish standard voice Mia. We increased the synthetic voice’s pitch to produce a teenager’s speech. We designed the robot’s verbal and non-verbal behaviors to guide and keep children engaged during the interaction. For the non-verbal behaviors, we manipulated the robot’s facial expressions to exhibit emotions (e.g., happiness, surprise, wink) and generate head movements (i.e., nodding and shaking). The verbal behaviors consisted of different utterances to have a general conversation (e.g., “Do you like to tell stories to your friends?”, “My lab is in a very cold city, how is the weather where you live?”), provide support (e.g., “Do you want me to help you?”), follow up on the story (e.g., “And what were they doing?”, “That’s a good idea”) and produce story ideas (e.g., “Once upon a time, there was an alien robot destroying the world”, “let’s say that there was a magic chicken that had a very good sense of smell”). The story utterances used by the robot were extracted from stories told by participants in our previous work [30]. We used a corpus of nine statements for each session, and these statements were different between sessions. Following good practices for user studies in HRI, the behavior of the robot was teleoperated by following a pre-defined interaction script.

C. Participants

The number of participants needed to detect an effect with 95% power and $\alpha = 0.05$ was calculated a priori using the Python library TTestPower for paired sample t-test, resulting in a sample size of 22. Therefore, a total of 25 children between 9– and 12–years old ($M = 10.3, SD = 0.8$) took part in the study. We recruited children from the fourth and fifth grades at a private school in Bogota, Colombia. Consent forms were sent to legal guardians, and only children whose informed consent form was signed and returned to the school participated in the study. Each child played two sessions of the storytelling game with the virtual Furhat. We excluded data from 7 participants for various reasons, such as stopping the activity prematurely, participating in only one session, or technical issues. After exclusion, we gathered data from 18 children (10 girls and 8 boys) to analyze the results. This study had Ethical clearance from the county’s regional ethics committee (Regionala etikprövningssämnand i Uppsala, document no. 2018-503).

D. Measures

This session presents the measures of this study, obtained by combining two data collection methods: questionnaires and objective measures based on linguistic behaviors.

---

2Web engine used for the graphical interface in this study: [https://playcanvas.com](https://playcanvas.com)

3Voice used in this study: [https://aws.amazon.com/polly/](https://aws.amazon.com/polly/)

---

4The materials of this submission were stored in the Open Science Framework (OSF) and can be found in the following link: [https://osf.io/vj53p/?view_only=c55b7f461b2545cab0eb476bb92eba4a](https://osf.io/vj53p/?view_only=c55b7f461b2545cab0eb476bb92eba4a)
Likability component of the Godspeed questionnaire [34]. Likability was measured using a modified version of the 5-items scale used in [33]. Competency trust was measured using a 14-items post-session survey on a 5-point Likert scale developed in [32]. Social trust was captured using the 4-items validated items for moral trust and six items for performance trust in [19] for moral, and performance trust. We selected four items for competency trust and four items for likability in the virtual Furhat robot, we implemented a 14-items post-session survey on a 5-point Likert scale. Social trust was captured using the 4-items validated scale developed in [32]. Competency trust was measured using a modified version of the 5-items scale used in [33]. Likability was measured using a modified version of the “Likability” component of the Godspeed questionnaire [34].

1) Moral and Performance Trust Beliefs (MPTB): To examine how children conceptualize trust and its relationship with trusting robots, we designed a short questionnaire following the Multi-Dimensional Measure of Trust proposed in [19] for moral, and performance trust. We selected four items for moral trust and six items for performance trust (Table I). Previous work has shown that children’s judgement of trust in robots is mainly driven by their friendships [31]. Hence, we asked children to remember how they interact with their friends and how they think it would be with general robots. They rated their beliefs on a 3-point scale (i.e., Good/Bad). We provided this questionnaire as a pre-test for both friends and robots, and as a post-test for robots, in the two interactive sessions.

2) Social and Competency Trust, and Likability (SCT&L): To assess children’s perceptions of Social Trust, Competency Trust, and Likability in the virtual Furhat robot, we implemented a 14-items post-session survey on a 5-point Likert scale. Social trust was captured using the 4-items validated scale developed in [32]. Competency trust was measured using a modified version of the 5-items scale used in [33]. Likability was measured using a modified version of the “Likability” component of the Godspeed questionnaire [34].

3) Self-Disclosure Task (SDT): Previous literature has shown that children disclose personal information to their friends, and the amount of information increments as their friendship becomes closer [35]. As such, increased disclosure of information produces trust [22]. We implemented a modified version of the Self-Disclosure Task developed in [25]. In this task, the robot disclosed two pieces of information (i.e., good and bad information about itself) and prompted for disclosure in return. The SDT happened at the end of each interactive session. The content of the information disclosed by the robot was different between the interactive sessions.

E. Procedure

A few weeks before the data collection, we introduced children to the study and the virtual robot at a class level in a video call. We allowed children to ask questions, but robot-related questions were postponed until the debriefing to prevent any influence on children’s perceptions of the robot. The experimenter emphasized that children’s participation was voluntary, that recorded data would be anonymized, and that they could withdraw from the study at any time without giving reasons.

The study was conducted for two consecutive weeks and was entirely online. We used the video conferencing tool Zoom. Due to the local COVID-19 restrictions, the school had a hybrid educational format. Therefore, some children were in their homes, and some were at the school (3 at home, 22 at school). The local school teachers assigned an unused room exclusively for the study, where a computer with a camera and microphone was placed. Two teachers and two students from the last grade were responsible for the coordination and logistics at the school. Their tasks included taking each participant to the room, starting the video call, and launching the game’s graphical interface.

Each child played a total of two storytelling sessions with the virtual Furhat robot—one session per week. The procedure was the same for each interactive session: at the scheduled time, the child was guided to the experimental room, where s/he was alone attending the video call. In the zoom room, three users were present: the experimenter, the participant, and a virtual user for the virtual robot (see Figure 1). The experimenter was responsible for explaining the activity’s objective to the participants, guiding them through the different stages, and teleoperating the robot. Before the interaction began, the experimenter asked for children’s verbal consent to participate in the study. Then the experimenter provided the first pre-test questionnaire. Once participants finished the first round of questionnaires, the experimenter introduced the virtual robot and the storytelling activity. Participants were told that they could use the interface to create the story they wanted, also that the robot would deliver some story ideas, but they were the leaders of the story. The experimenter turned off her camera to avoid participants getting distracted by her presence. Once the participant finished the story, the experimenter turned on her camera and provided the second round of questionnaires. The camera of the robot was turned off when participants were answering questionnaires. Later, the experimenter introduced the final activity, where participants played the SDT with the robot. Once participants completed the final activity, they were debriefed. The experimenter answered participants’ questions and explained the functionality of the virtual robot.

Since children did not have a time limit in the interaction with the virtual robot, the total duration of the study ranged between 20 - 35 minutes per child. The study was conducted entirely in Spanish.

V. Results

All analyses were conducted using Python 3.8. We tested our questionnaires to measure internal consistency reliability using Cronbach’s alpha. The trust’s beliefs questionnaire (MPTB) presented high internal consistency for moral trust Cronbach’s $\alpha = 0.72$ and performance trust Cronbach’s
The results of the Wilcoxon signed-rank tests revealed a non significant effect of the Interactive Session on the total number of words used by children in the SDT. We counted the number of responses, the number of spoken words after the robot disclosed information, and the number of spoken words after the robot’s prompts. We also analyzed the content of the information disclosed. Following the procedure proposed in [25], we classified children’s responses content into six categories presented in Table III. We included the category “Empathy” to analyze whether children responded empathically toward the robot’s disclosure.

The results of the Wilcoxon signed-rank tests revealed a non significant effect of Interactive Session on the number total words (W = 47.0, p = 0.09), responses made (W = 25.0, p = 0.26), and number of words after robot’s disclosure (W = 48.5, p = 0.51), see Figures 5(a), 5(b), and 5(c). A statistical effect of Interactive Session was found on the total number of words after the robot’s prompt (W = 31.5, p = 0.018). Children significantly disclosed more information after robot’s prompt in S2 (M = 25.55, SD = 13.75) compared to S1 (M = 17.55, SD = 16.12), see Figure 5(d).

We qualitatively analyzed our data regarding the content of the information disclosed. We observed that in general children disclosed more information related to the categories Physical and Fine Motor skills, and that this the amount of information was higher in S2 (M = 1.38, SD = 1.28; M = 1.83, SD = 1.46) than in S1 (M = 1.11, SD = 1.28; M = 1.30).

### Table II: Means and SD’s: Children’s Trust Beliefs

<table>
<thead>
<tr>
<th>Trust Dimension</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Friends pre-test</td>
</tr>
<tr>
<td>Moral</td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>1.69 (0.38)</td>
</tr>
<tr>
<td>Session 2</td>
<td>1.76 (0.29)</td>
</tr>
<tr>
<td>Performance</td>
<td></td>
</tr>
<tr>
<td>Session 1</td>
<td>1.69 (0.24)</td>
</tr>
<tr>
<td>Session 2</td>
<td>1.71 (0.30)</td>
</tr>
</tbody>
</table>

\( \alpha = 0.82 \). Our second questionnaire (SCT&L) also presented high internal consistency for the three constructs: social trust Cronbach’s \( \alpha = 0.84 \), competency trust Cronbach’s \( \alpha = 0.69 \), and likability Cronbach’s \( \alpha = 0.93 \). We performed the Shapiro-Wilk test for normality. The null hypothesis was rejected (\( p < 0.05 \)) for all our variables, indicating that our data deviated from a normal distribution. We used the Wilcoxon signed-rank test, the non-parametric version of the paired t-test.

### A. Children’s Trust Beliefs

RQ1 is dedicated to examine (1) whether children’s conceptualization of trust is multidimensional, (2) whether children’s trust beliefs in friends and general robots remain constant over time, and (3) whether children’s trust in robots is similar to the way they trust in their friends. We coded children’s responses to the MPTB questions on a 3-point scale (i.e., No = 0, Maybe = 1, and Yes = 2). The results of the Wilcoxon signed-rank tests revealed a non significant effect of Interactive Session on moral (\( W = 14.0, p = 0.29 \)) and performance (\( W = 53.0, p = 0.69 \)) trust in friends, and on moral (\( pre : W = 12.0, p = 0.73; post : W = 13.5, p = 0.07 \)) and performance (\( pre : W = 13.5, p = 0.15; post : W = 11.5, p = 0.36 \)) trust in general robots. Our results revealed a non significant effect of the stage of the interaction (i.e., pre-test/post-test) on children’s trust beliefs in general robots for moral trust (\( S1 : W = 41.0, p = 0.46; S2 : W = 15.5, p = 0.72 \)) and performance trust (\( S1 : W = 32.5, p = 0.96; S2 : W = 16.5, p = 0.47 \)). Finally, we found a non significant effect of the type of agent (i.e., friends/robots) on children’s moral trust (\( S1 : W = 31.0, p = 0.52; S2 : W = 13.5, p = 0.28 \)) and performance trust (\( S1 : W = 64.0, p = 0.83; S2 : W = 47.5, p = 0.28 \)) trust beliefs. See Table II for means and SD’s.

### B. Social and Competency Trust, and Likability

RQ2 aims to investigate children’s perceptions of the virtual Furhat robot by comparing scores of Social Trust, Competency Trust and Likability (SCT&L) between S1 and S2. Results showed a significant effect of Interactive Session on Competency Trust (\( W = 29.0, p = 0.04 \)), children rated the robot as more competent and reliable in S2 (\( M = 4.51, SD = 0.51 \)) compared to S1 (\( M = 4.33, SD = 0.45 \)).
TABLE III
CATEGORIES FOR THE TYPE OF INFORMATION DISCLOSED

<table>
<thead>
<tr>
<th>Category</th>
<th>Samples of children’s responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical Skills</td>
<td>“I can ride a horse”</td>
</tr>
<tr>
<td></td>
<td>“I can’t run very fast”</td>
</tr>
<tr>
<td>Fine Motor Skills</td>
<td>“I’m good at drawing”</td>
</tr>
<tr>
<td></td>
<td>“I can’t play instruments”</td>
</tr>
<tr>
<td>Cognitive Skills</td>
<td>“I think I am good at math”</td>
</tr>
<tr>
<td></td>
<td>“I am not very smart”</td>
</tr>
<tr>
<td>Social Skills</td>
<td>“I am a good friend”</td>
</tr>
<tr>
<td></td>
<td>“I don’t know how to help others”</td>
</tr>
<tr>
<td>Non Specific Skills</td>
<td>“I am good at many things”</td>
</tr>
<tr>
<td></td>
<td>“I am not perfect”</td>
</tr>
<tr>
<td>Empathy</td>
<td>“I am sure you will make it”</td>
</tr>
<tr>
<td></td>
<td>“You should not be worried about it”</td>
</tr>
</tbody>
</table>


VI. DISCUSSION

A. Trust Beliefs in Friends and Robots as a Baseline (RQ1)

The goal of this RQ was to investigate the essence of children’s concept of trust beliefs in friends and the relationship with trusting robots in general. We found that children can make judgments of moral trust and performance trust, which suggests that children at this stage of development can understand the multidimensional nature of trust. Our findings are in line with studies from social psychology that suggest that children at a very young age can evaluate competence and benevolence when choosing whom to trust [16]. Moreover, we did not find significant effects of Interactive Session on children’s trust beliefs in friends and general robots. This result indicates that their concept of trust remains constant over time. We found this particularly important as some studies suggest that trust is highly personalized and can change over time, but in middle childhood, children have temporal consistency of trust [18]. This result introduces a multidimensional “trust baseline” to children’s general tendency to perceive the trustworthiness of friends and general robots that holds a temporal consistency.

Fig. 5. Analysis of children’s disclosure of information per Session

However, these findings should be interpreted with caution. First, understanding the relationship between trust beliefs in friends and robots is complex and varies based on the individual. For example, earlier work has demonstrated that children form high expectations of the trusting relationship with friends [11] and these expectations can be similar when trusting robots [36]. Nevertheless, as children have more experience with robots and technology, these expectations might not be sustained, resulting in a change in their trust beliefs [8]. If this happens, the “trust baseline” needs to be adjusted. A second implication concerns the effect of time on children’s concept of trust. Our study assessed children’s trust beliefs over two weeks, yet we can not argue whether this behavior will hold for a more extended period. To further investigate this, longitudinal studies need to be conducted over several weeks to assess children’s trust beliefs in friends and robots over time.

B. The Effect of Time on Trust Development (RQ2)

We found that the second encounter with the robot increased children’s perceived competency trust in the virtual robot. One possible explanation for this result is that in S2, children were generally more familiar with the robot’s
behavior and the storytelling game. We believe that this could lead children to exploit the robot’s capabilities further (e.g., asking the robot to contribute to the story plot or integrating the robot’s ideas into their stories) and hence have a better performance at storytelling in S2. Previous studies have found that several encounters can benefit language learning with robots, especially when the robot’s behavior matches the children’s language ability [2]. In our study, the robot’s role was to help children create stories, and the robot’s ideas matched the scenarios and characters displayed in the graphical interface. Therefore, in this work, we believe that in S2, children were more acquainted with the game and the robot’s role and felt that their performance increased compared to S1 resulting in higher competency trust in the robot. To provide evidence to this assertion, in future work, we plan to analyze children’s storytelling creation process and the effect of Interactive Session and robot behaviors on children’s verbal behavior (e.g., Do children include the robot’s ideas in their stories more often in the second session?).

Nevertheless, the same effect was not found on social trust and likability. We observed that there were ceiling effects for both constructs in both sessions, which is not surprising for subjective measures that focus on the experience of the interaction. Our results align with previous studies that found that children’s perceptions of robots remain positive over time [1], [3]. Thus, we believe that the novelty effect was still present in the second session, and children continued having higher expectations of the robot. Future work should investigate how the novelty effect affects trust dynamics by looking at short-encounters and long-term interactions.

C. The Effect of Time on Relationship Formation (RQ3)

Children significantly disclosed more information to the robot’s prompt in S2. This result suggests that children could have felt closer to the robot in the second interaction and were more likely to reveal more personal information. We presume that children perceived the robot as more social trustworthy over time as sharing personal information about oneself is a crucial component of close relationships [23]. Furthermore, children disclosed more information about physical and motor fine skills in S2, and their responses were more elaborated than in S1. This observation might indicate children’s perception of their competence and how this can affect their relationships with others. Evidence suggests that during middle childhood, children provide more information about their competence due to their concern at gaining the initial peer acceptance [37]. Although this behavior could happen in our study, we can not draw strong conclusions from our data. This will be an interesting line for future research to assess individual differences and their implications in trustworthy relationships with social robots. The results showed that time elicited more disclosive behaviors concerning social skills and more empathetic responses toward the robot in S2 than in S1. The robot’s disclosure could foster closeness, which could be reinforced by the second encounter, positively influencing the relationship formation.

D. Ethical Considerations in cHRI

There are ethical concerns regarding deception when designing social robots that aim to build relationships with children, especially when investigating trust. First, the robot was operated using the Wizard-of-Oz technique; this might affect children’s perceptions of the robot as shown in [38]. In this case, it is crucial to consider the system’s transparency and inform users that a human controlled the robot. In our study, we did not further investigate children’s perceptions after debriefing, which could be interesting for future research that aims to mitigate deception and encourage good ethical practices in cHRI. Second, during the SDT, disclosing information was prompted by the robot revealing its physical and cognitive capabilities. This is fairly realistic as robotic systems can be integrated with these features, but introducing robots as agents with particular interests or preferences should be taken with caution to avoid false expectations.

VII. LIMITATIONS

COVID-19 introduced challenges when conducting user studies in cHRI. On the one hand, we could test our application with broader populations making social robots more accessible to different cultures. On the other hand, we were limited to online settings. This was a limitation in our study as we can not know yet if our results will hold when participants interact with the physical Furhat robot. In future work, we may replicate the study to investigate whether playing the storytelling game with the physical robotic embodiment will yield different results in children’s trust in the social robot.

In our current study, a more detailed investigation of the effect of robot behaviors on children’s performance and its relationship with trust development is lacking. We collected video data to facilitate this assessment in the future.

Our study showed that the novelty effect can still be present after seven days. To investigate how and when the novelty effect disappears and whether it influences children’s trust development with robots. Future work should also account for several encounters between children and the robot over extended periods.

VIII. CONCLUSION

In this work, children interacted with the virtual Furhat robot in an online storytelling game in two different sessions with seven days of zero exposure in between. We found that children’s multidimensional trust beliefs in friends and general robots remain constant over time and could be set as a “trust baseline”. Our work also showed empirical evidence using subjective and objective measures that time is crucial for children’s trust development with social robots and their relationships with them. Specifically, the perception of competency trust, and closeness of the virtual robot was higher in the second encounter. Our results stand out as they show that repeated encounters might impact children’s trait inference over time. This is particularly important when assessing the long-term acceptance of social robots used as companions in children’s education.