UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering, Mathematics and Computer Science

Agents presenting themselves as Strangers during Privacy Permission Requests: Effects on Disclosure and Privacy Awareness of Children

Thesis MSc Interaction Technology

Author:
Nynke Zwart

Supervisor:
Khiet Truong

January 31, 2021

Contents

Abstract

Young children are prone to create strong bonds with robots and are very trusting of companies up until the age of 12. They give out a lot of information about themselves, even though they seem privacy aware. In recent papers, it became clear that the embodiment of a robot can facilitate higher levels of information disclosure, which creates serious privacy concerns as no regulations exist on this. Furthermore, not much research about this phenomenon has been done in relation to children, even though children are in contact with robots from a very young age and create strong bonds with them: this can make them more vulnerable to disclose information. These gaps in knowledge led to the experimental study (a 2x2 between-subject design) of this report in which 79 children, aged 8-12 years old, conversed with an embodied conversational agent for five minutes. The agent was either a Furhat robot or a Google Home Mini device. At several points in the conversation, the robot requested personal information in the form of a privacy permission request. Children's compliance on this determined an information disclosure score and their understanding of the content of the request determined a privacy awareness score. Besides the different levels of agent embodiment, a "stranger presence" within the agent could occur during a request. For the Furhat, this meant a change in voice and face, for the Google Home Mini a change in voice only. The stranger symbolised the company behind each service, as the people behind a company are often strangers to its users. Next to this, children are taught about not complying to strangers from a young age and they indicate that they seek privacy from strangers online as well. This novel approach of a stranger presence might therefore trigger their knowledge on stranger danger and lower their information disclosure. The results show that the level of agent embodiment and presence or absence of a stranger did not lead to significant differences in information disclosure and privacy awareness. Observations, children's comments and post-questionnaire measurements are used to go into further depth on these results.

Keywords

Child-Robot Interaction, Stranger Danger, Embodiment, Information Disclosure, Privacy Awareness, Conversational Agents

Project website

`https://sites.google.com/view/master-thesis-nynke-zwart/`

1 Introduction

Privacy is often defined as "having control over information about ourselves" [34]. In today's society, it can feel overwhelming to achieve this, as every website, robot or other service has its own privacy policy that is filled with jargon and lengthy sentences [3], [35], [46], [57], [61], [70], [76]. As a result, people become either more concerned about their privacy [70] or they simply accept the privacy policy without reading it [33], [95]. It is not uncommon for people to think they are privacy aware and still hand out much information about themselves at the same time. A privacy paradox can be recognised in this: a discrepancy in people's intentions and actual behaviour [67]. Many researchers have looked at different ways of changing privacy policies to increase privacy awareness, which is often measured by policy understanding [19], [26], [46], [77]. A promising method makes use of timing [26], to request information from people when this information becomes relevant to collect. This way, users do not have to agree with everything beforehand. A privacy permission request contain small parts of a privacy policy, which are easier to process for people than the full document. Most research on privacy has been done with adults, but children are an important group to do privacy research with as well, because they are also subjected to the privacy paradox.

It seems that children are quite privacy aware, as has been found through mainly qualitative research, such as interviews and co-design sessions [53]. They have different strategies to keep themselves safe online [5], [20], [50], [53]. These can be strategies that are implemented by their parents but also ones that they have come up by themselves, such as falsifying information and changing profile settings to “friends only”. However, many do not read privacy policies, as this is already complicated to adults. They also find it hard to explain possible consequences of online risks [96]. Furthermore, children still give out a lot of information about themselves online, with children (aged 14 and below) disclosing more information about themselves than their older peers [8], [21], [49]. A privacy paradox seems to exist here as well. According to EUKidsOnline, that published a ranking of the biggest online risks for a child in Europe, giving out personal information was found to be the biggest risk [40]. It is a risk because misuse of personal information can occur in several ways, for example through cyberbullying, identity theft and data abuse by companies. Therefore, it remains relevant to research in which contexts and why this paradox occurs, especially because children go online a lot and are allowed to be online, unsupervised, from as young as 8 years old [12]. They get their first phone at age 10 [17] and many children own multiple devices, such as gaming devices, tablets and computers. Next to this, embodied conversational agents such as Alexa and Google Home are found increasingly in people’s homes, with children interacting with them from as young as 15 months old [74]. Children thus have to learn how to navigate the “digital” landscape from a young age.

Moreover, recent research has found that that embodied agents can facilitate higher information disclosure, which is a cause for concern [10], [84], [86]. Similar results are found when a person has physical contact with an agent and when the agent has a high level of sociability [73], [75]. Most companies that create social, embodied agents use privacy policies that are similar to those for websites and apps. Current rules and regulations namely do not take the seeming advantages of embodiment into account [10], [42], [29]. Besides, most research on the topic of disclosure and agent embodiment is done with adults, which means there is a knowledge gap when it comes to children. It is known from research that children especially can create very strong bonds with robots [44], [45], [60], [74]. They could be even more vulnerable than adults when it comes to information disclosure to embodied entities [48], [85]. Once again, it is important to take this group into account.

It is because of the above mentioned reasons that this report focused on children (aged 8-12 years old) and created an experiment with different levels of embodied conversational agents: a Google Home Mini and a Furhat. Depending on the condition, the embodied agent held a conversation with a child in which privacy permission requests occurred, to obtain privacy sensitive information. Such privacy permission requests were used to determine a child’s privacy awareness. As for information disclosure, the embodiment of the conversational agents afforded the introduction of a novel method: namely, turning the agent into a stranger.

Children are very aware that they should not comply to requests from strangers, to avoid ”stranger danger”. From a young age, children are taught about this - the risks that complying to an unknown person can bring - by their parents and educators [6], [14], [20], [52], [59], [65], [4]. When asked whom they seek privacy from online, 79% of children answered “from strangers” [20]. This phenomenon can be interesting to apply to entities that wish to collect personal information, to find out if it can lessen information disclosure of children. This would then introduce a novel approach, that tries to leverage knowledge on stranger danger. Furthermore, a metaphor can be drawn between a stranger appearance and a company. A stranger appearance can represent the company behind the agent, which is run by people that are often strangers to their users. In general, children are very trusting of companies up until the age of 12 [53]. Perhaps this is the case because companies are not visible/graspable and children prefer to frame things in a real world context [53]. A stranger presence within a conversational agent could be a first step towards raising awareness on this. That a different presence might heighten awareness, has also been proposed in a research on Alexa and third-party apps [56]. It was shown that many Alexa users did not recognise whom they were talking to at times and freely gave away information to third party apps without knowing. The researchers suggested a change in voice might heighten users' awareness. This inspired the agent "changes" within this report as well. In terms of the agents used in this report, the Furhat robot afforded a change in voice and face and the Google Home Mini afforded a change in voice only.

Taking all these elements into account, a novel way of researching information disclosure and privacy awareness of children was thus tested. This was done using different embodied conversational agents that could facilitate a stranger presence within an agent. Each participant conversed with an embodied conversational agent that wished to collect personal information at various points in a conversation. The aim of the research was to study the influence of embodied conversational agents, as well as the absence or presence of a stranger within this agent during privacy permission requests, on information disclosure and privacy awareness of children. It was hypothesised that children would give away less information to a stranger presence within the agent, as opposed to when the agent stayed the same during the entire conversation, mainly because they are taught to not comply to strangers. The higher embodied agent was thought to match children's expectations of a stranger more, leading to a bigger impact. Therefore, it was hypothesised that less information would be disclosed to a stranger presence in a high embodied agent than a low embodied agent. When the agent stayed the same during the entire conversation, the opposite effect was expected. As for privacy awareness, it was hypothesised that children would be more privacy aware in the stranger presence conditions, because the stranger would facilitate more attention from them in order to minimise the risk of stranger danger, which would make them remember better what was said during the privacy permission requests. A higher embodied agent was also hypothesised to lead to higher privacy awareness scores. More detailed explanations are written in the hypotheses section of this report.

The rest of the structure of this report is as follows: it starts with a literature study about the related work on this topic, followed by the research question and hypotheses. Next, two pilot studies are described before the main study comes into focus. All studies include results, a discussion and limitations/future work. Lastly, an overall conclusion is drawn, which closes the report.

2 Literature study

2.1 Privacy policies

Privacy plays a large role in this research, so the literature study will first dive into this topic, focusing mainly on 1) the rules and regulations surrounding privacy policies, 2) the understanding of users of these policies and 3) the representation of privacy policies. Most research by far has been done in the context of websites or applications: albeit not directly connected to conversational agents, as used in this report, the section is useful for several reasons. The first mentioned point is relevant, in order to understand what is currently being asked of companies to provide to users. The second point contains research that shows how adult users struggle understanding privacy policy documents. It becomes clear that it cannot be expected from children to read these documents to gain privacy awareness. Point 3 contains research that has tried out different approaches to heighten privacy awareness through different privacy policy representations, which inspired elements that are used in the main study.

2.1.1 The definition of privacy

The term privacy has many definitions, such as “the right to be let alone” [88], “the condition of being protected from unwanted access by others” [13] and “having control over information about ourselves” [34]. A definition by Alan Westin, a pioneer in the field of privacy, is often quoted. His definition states that privacy is an individual’s control over information that is knowingly given or shared with others [90]. This also includes when information is obtained and what uses will be made of it by others. The latter grows more important by the day as companies collect huge amounts of user data. All kinds of information are saved and even shared between companies and other parties online. The commercial value of personal data keeps on growing. The privacy guidelines of Google are an example of this shift in business model. In 1999, its first privacy policy stated that Google would only share “aggregate” information to advertisers, business partners, sponsors and other third parties: “We only talk about our users in aggregate, not as individuals” [39]. For example, information about the number of visits to Google per day would be only one number, aggregated over all its users. Over the years, this has changed to the following: “We will share personal information with companies, organisations or individuals outside of Google when we have your consent to do so.” Personal info can include search queries, clicks, uploaded content and more. Whereas one user profile might not be very profitable, the real value comes from huge numbers of data profiles. Big tech companies offer targeted marketing based on the data provided by its users. The services of the company may thus be free, but are in essence “paid” for by the user, through giving away their data. Therefore, the second part of Westin’s definition is something to be cautious about as a user. There are ways to inform ourselves on company practices, for example through privacy policies. Such documents are made to inform users about data collection practises of companies and can help to raise privacy awareness. However, it will become clear in the next section that these documents are not easy to get through.

2.1.2 Perception of privacy documents

The previous paragraphs briefly touched upon a few lines from Google’s privacy policy. A privacy policy is a legal document that contains information about how a party gathers, uses, discloses and manages a customer’s or client’s data. The term is used interchangeably with “privacy notice” and “privacy guidelines”.

A privacy policy can be found on nearly every website and application. There is no worldwide set of rules that companies need to comply to: each country or continent has their own laws. However, these rules cross borders. This means that an American website needs to fulfil the EU law requirements whenever they collect personal information from a European user or transfer personal information to- and from an EU country. This makes privacy policies complicated to create and maintain as so many different requirements are asked of companies. A few of the most well known laws are:

- **CalOPPA**: The California Online Privacy Protection Act, effective since 2004, is the first state law in the US that requires companies to include privacy policies on their websites [42]. The policy must include the information that can be gathered by the website, how the information can be shared with other parties and - if possible - how the user can change or review the stored information. The CalOPPA made it more common to publish privacy policies on US websites, since it is highly likely that somebody from California will visit a US website. Because the CalOPPA is stricter than other rules in the US, websites will sometimes have a separate section named “Your California Privacy Rights”.

- **GDPR**: Standing for General Data Protection Regulation, this regulation is created to provide EU and European Economic Area users more control over their personal data. The law went into effect in 2018, taking over the Data Protection Directive that was adopted in 1995. A big change that was introduced is having the right as a user to ask for a free detailed copy of all the data that is collected about them. They can also demand to have all their data deleted for good. Another requirement of the GDPR is that privacy polices must be written in “clear and plain language” that is “concise, transparent, intelligible and easily accessible” (Article 12) [29]. Furthermore, when privacy sensitive information is to be collected, the user has to give consent and be informed about the why’s and how’s accordingly (Article 7) [30]. They must be able to easily withdraw their consent at any given time.

Privacy Policies are written to inform the client how they handle their gathered data and what they use it for. However, studies with adults, before the introduction of the GDPR, show that users struggle with these legal documents, due to the lack of clear and plain language. The sentences and words that are used in many documents, are too long. Technical jargon is also often not understood by adults. Various methods that measure people’s understanding of texts, such as the Flesch Reading Ease Score and the cloze test [79], confirm this [3], [58], [76]. Milne found that if a privacy policy is not perceived as comprehensible, it is less likely to be read [61]. Although privacy policies should be written to protect a client, Earp et al. observe that they serve the company more than the client in protecting them from potential privacy lawsuits [25].

Another problem caused by companies is that their policies are often not straightforward in their practices. Privacy concerns of users are rather intensified than clarified because it is unclear to what extent companies use their data [70]. Furey and Blue’s research adds to this that the user is only vaguely informed about what other parties can access their data and what data exactly [35]. This lack of transparency and accountability does not improve the understanding of policies.

Thirdly, some websites are designed to make it harder for users to find the policy. When a privacy policy is not shown immediately, users avoid reading it [78], [92]. Those who do find and click on it merely scan through it. Most people however, never click on it. Steinfield found these results using eye-tracking experiments. Moreover, users base a company’s credibility on the ease of use and feel of a website, which can lead to false assumptions about privacy protection [33]. Even if people see security warnings in the toolbar, they often explain away warnings because they trust how the content on the website is presented [95].

Much research has been done in order to tackle the readability issues and make policies more appealing to read for users. This is important, as a good understanding of a company’s practises can raise users’ privacy awareness. The next section will go into detail on the research on this topic.

2.1.3 Different privacy policy presentations

This section will cover a collection of studies that introduce ways to change the presentation of privacy policies, in order to improve users’ privacy awareness. There are different approaches to measuring privacy awareness: the approach that is often used in the studies mentioned in this section and that can be checked for validity, is answering questions on understanding and memorisation of privacy policy information [11], [19], [46], [58], [72], [77].

Changing textual presentation Researchers have tried to create standardised forms for privacy policies, such as layered formats that require clicking to get into more depth or bulleted policy formats. These standardised forms increased reading speed, but still did not perform well on comprehension and joy in reading them [58]. Another study tried out a textual grid, but it missed a focal point and was found to be too cluttered, thus not performing well on understanding [72].

Changing visual presentation Kelley et al. created an improved grid layout by introducing visual labels that were inspired by nutrition labels and labels of energy companies [46]. This lessened the amount of text in the grid, making the layout clearer. Participants' understanding scores were higher for this presentation than a "normal" policy. Other researchers tried to develop even more compact visualisations, which were fully icon-based or used sentences that contained icons and relationship-arrows [71], [38]. Both were not evaluated by users. Lastly, Soumelidou and Tsohou created "Tag Clouds" to explain sections of privacy policies. Users' showed higher privacy awareness levels for the cloud method, compared a normal privacy policy [77].

Changing timing of presentation Research showed that if a policy is not presented immediately, it is less likely to be read [78]. Timing is thus an interesting component that can influence privacy awareness. Egelman et al. conducted a study on the placement of privacy indicators and the timing of their placements [26]. It turned out that timing had a significant impact on how much extra money participants were willing to pay for a safer site. Participants that were in conditions where they could immediately scan privacy ratings of websites also made fewer other searches and were quicker in their buying decisions.

Presentations in the "wild" Despite the above mentioned research, most websites still use traditional non-standardised text-based policies, which are not beneficial to users' privacy awareness. Since the introduction of GDPR, companies such as Google have added short animated videos to their privacy policies. However, no research has been done on whether this is beneficial to users. Another approach that is becoming more popular, which depends on timing and relevancy, is the pop-up permission request. These requests are often used in applications and websites. A privacy permission request contains a small piece of information from a privacy policy (called context), that is related to the information to be collected. It informs the user why and how this information will be used. Privacy permission requests usually only become visible when it is relevant to collect the requested information. Timing is thus an interesting component to take into accountin the study of this report, especially considering Egelman et al.'s positive results on people's privacy awareness.

As mentioned previously, this section only discussed research with adult participants, while the research group of this report is children. It may be derived from reading this section that it cannot be expected of children to understand a full privacy policy, when adults already struggle with these documents. Therefore, other representations should be looked at that can help to raise children's privacy awareness. The above mentioned permission requests are perhaps a suitable approach, as they offer smaller chunks of information to process, instead of a full document. The next sections will go into more detail on the research on children and privacy that has been done previously.

2.2 Children and privacy

This section explains why privacy policy understanding is relevant to study with children. It starts with describing the current situation and continues to discuss research on how children perceive privacy risks and the privacy strategies they use. Further insight is then provided on children's privacy awareness and information disclosure and the privacy paradox that is connected to this. Furthermore, "stranger danger" is discussed and how children seek privacy from strangers online.

2.2.1 Children on the internet

In 2002, 90% of US children between 5 and 17 years old went online, more than any other age group. For 10 to 13 year olds specifically, this percentage was 65% [69]. To add to these facts, more than 175,000 children worldwide go online for the first time every day, according to a press release in 2018 by UNICEF [81]. The average age of these children is unknown, but from earlier surveys as done by Microsoft in 2013, it was found from 1000 respondents (parents and non-parents) that on average, they would allow their children to browse independently on devices and on the internet from the age of 8 [12]. Of all parents with children under the age of seven, 41% allowed unsupervised access to a computer and 29% allowed unsupervised mobile app use. As these numbers are from 2013, it can be expected that they are already higher than this. The size of children's own device collection most likely contributes to this increase. In 2016, 64% of children accessed the internet on their own laptop and tablet, which was only 42% in 2012. Furthermore, the average age of getting a first phone was found to be 10.3 years old [17]. This questionnaire was filled in by 500 US parents.

In Europe, Dutch children are the highest users of internet (93%). This was found through a survey by EU Kids Online, which was based on parents and guardians with children less than 18 years [40]. They fall into high-risk categories when it comes to safety and privacy, especially because children use public applications (such as Youtube, Snapchat and Instagram) more than applications aimed at children [53].

It is therefore important to do privacy research with children, not only because they are young, but also because they are often unsupervised users and fall into these high-risk categories. The next section will first go into detail on the rules and regulations that are set in place to protect children's privacy.

2.2.2 COPPA and GDPR-K

A US Federal Trade Commission survey of 212 children's websites in 1998 found that 88% of all sites collected at least one type of personal identifying information (name, email-address, postal address, age, hobbies and others) [18]. 21% of sites collected five or more types of information. 46% of sites did not have a privacy policy or information statement or both. Rules and regulations were set in place in the 90s to restrict these practises, specifically COPPA. The GDPR-K is also a well known, recently introduced law. This is what they entail, in short:

- **COPPA**: COPPA stands for Children's Online Privacy Protection Act, which went into effect in 1998 in the United States. COPPA changed the collecting practices for children under thirteen years old: websites and online services need to conform to special rules. For instance, there must be a clear privacy policy about how a child's personal data is handled. Furthermore, parents need to give additional consent about collection of their child's data and have rights to change privacy settings from little to no collection and to request for all data to be deleted. Companies also need to create extra security, confidentiality and integrity around children's data [1].

- **GDPR-K** The EU General Data Protection Regulation contains sections about children's (Kids') data collection, which are referred to as GDPR-K. Article 12, that requires clear language in policies, emphasises this needs to be taken into account when "any information [is] addressed specifically to a child". Article 8 governs the process of obtaining consent from parents for their child and the validity of this consent [31]. The GDPR-K specifically mentions that children deserve special protection "as they may be less aware of the risks, consequences and safeguards concerned and their rights in relation to the processing of personal data" [31]. A difference between the GDPR-K and COPPA is that member states in the EU can decide by themselves what the maximum child age is, ranging from 13 to 16. By default, the chosen age is 16.

These regulations thus require extra steps to be taken by companies to align their data collection practises. As a result, there are companies that have set a 13 year age limit before an account can be created, such as YouTube (not to be confused with YouTube Kids), Snapchat and Instagram. As children themselves indicated that they use public applications [53], ways are found to circumvent these rules. It is also known that children go online unsupervised from the age of 8, so this means that children aged 8-12 are in a grey area of self-regulation and being protected by rules. This is the reason why this age-group is studied in this report.

The next sections explore the situation of children further, in terms of risks and parental and children's privacy strategies.

2.2.3 How parents view children's privacy

In 1999, Ackerman et al. created a survey on privacy in e-commerce and asked 381 adults to fill in a survey on what pieces of information they were comfortable giving out to websites. This information could be an email address, age, a favourite snack, and more. Ackerman also asked users' comfort in giving out this information if it would concern a child in their care between the ages of 8 and 12. Participants were significantly less comfortable with this on all questions [2]. Whereas Ackerman's findings are very specific, most research focuses on bigger concepts, such as fear. In boyd's and Hargittai's research, common fears were rated by parents with 10-14 year old children (N=1007), such as their child meeting a stranger they know from the internet, being exposed to pornographic content, being exposed to violent content, being a victim of online bullying and bullying another child online [14]. The levels of concern were expressed in this order, with meeting a stranger being the highest fear: 63% of parents were extremely concerned. 30% of EU children (N=25000) between 9 and 16 years old has had contact online with someone not met face to face [51], but the percentage of parents whose children actually experienced harm by a stranger was 1% [14]. In reality, if harm does occur, it usually does not involve strangers, but people known to the victim [62]. It is not unusual that risk is badly assessed by humans, especially in fields where they are not experts. This badly assessed risk can be attributed to the media, that arguably have created the biggest fears around the topic of "stranger danger" [52]. However, since the awareness around this topic is high, it is interesting to look into this phenomenon in further detail (see Section 2.2.6).

Another research, by Madden et al., looked at concerns of parents with 12-17 year old children (N=802). Questions were asked about interaction with strangers, but also about information available to advertisers, impact on future opportunities and reputation management. 53% was very concerned about interaction with strangers, closely followed by what information about their children was available to advertisers (49%) [55]. Most social networking companies get their income through personalised advertisements [43], that are based on information disclosed by their users. The latter concern thus aligns better with the real world.

Parents' concerns about their children online lead to different strategies. During interviews, several strategies were mentioned. For example, some parents let their children use password protected devices that only they knew the password of. Others only put devices in the same room as where they were present, so they could monitor the child [41]. Only a few of the 14 participating parents in Zhang-Kennedy's research used more advanced strategies such as parental control tools on the device or changing privacy settings [96], compared to none of the 18 parents in Hsiao's research. In Hsiao's research, none of the participants were even aware of COPPA [41].

The parents in Madden et al.'s larger survey were more privacy aware. For example, 44% of all parents had read a privacy policy for a website or application that their child was using. According to the research findings, the level of education of the parent, as well as being a user themselves, heightens chances of having read a privacy policy [55]. Next to this, 31% of parents helped their child with privacy settings and 50% of parents in their research used parental controls. Most parental control apps prioritise monitoring and restriction instead of promoting communication and self-regulation amongst children [93]. Parental influence is important, as adolescents who value privacy in the real world are also more careful with disclosing private information online [21]. However, too much of it can destroy trust and respect between parent and child. Therefore, researchers advise that parents should leave room for the child to learn how to navigate the digital world on their own [66]. This report therefore focuses on a novel research method that does not include parents.

The next sections will dive into what children think about the topic of privacy, the ways they monitor themselves and what they expect of their parents, to get a sense of their privacy awareness.

2.2.4 How children view privacy risks

Livingstone et al. performed a large database research on 105 articles to map children's privacy awareness [53]. They mapped these findings in age categories of 5-7, 8-11 and 12-17 years old. Firstly, 5 to 7 year olds know how to use digital devices for a small set of activities. They understand the idea of secrets and how to hide them, but have a limited understanding of the risks in the digital world. They come up with risks that link to the physical world, such as protecting themselves from siblings who steal their devices. 8 to 11 year olds start to understand digital risks better. They can name online privacy risks such as "stranger danger". 5-11 year olds are in general very trusting of companies. Lastly, 12 to 17 year olds are more wary about trusting companies, gain more privacy understanding and are familiar with concepts such as data traces. However, they have little concern about possible future consequences. According to Zhang-Kennedy's research, children aged 7-11 find it hard to explain possible consequences of risks, as they are only aware of the basic concepts [96].

Another research by Livingstone et al., with 8543 children, reveals the risks that children come up with by themselves [52]. 55% of all the mentioned risks were content risks, with pornographic content being the highest risk (22% of all risks), followed by violence (18% of all risks). Contact risk, specifically "stranger danger", made up 14.0% of all named risks by the children. A contact risk "positions the child as a participant in adult-initiated activities, possibly unwillingly or unwittingly". The low percentage might result from the fact that it is a risk that is not very likely to occur, so children do not have much experience with it [14], [40]. It might also be due to the fact that children are taught less about stranger danger in online situations. Section 2.2.6 describes teaching methods in more detail to get insight into this.

Lastly, giving out personal information was named as 3% of all risks [52], while it

Online risks for children
Giving out personal information
Encountering pornography online
Seeing violent and hateful content
Being cyber-bullied
Receiving unwanted sexual comments
Meeting an online stranger offline

Table 1: This EUKidsOnline ranking is based on 390 empirical studies conducted in Europe that meet certain quality thresholds. It is ranked from most common to least common [40].

is actually the biggest online risk for children, according to EUKidsOnline [40]. Hence, this risk is important to look at into further depth, as will be done in section 2.2.7. The full ranking of risks can be found in Table 1. Next, the way children handle their privacy will be discussed.

2.2.5 How children handle their privacy

In general, parents try to handle their children's privacy, but from age 11 onwards, children showed privacy tactics in the focus groups of Livingstone et al. (N=135, ages 11-16). Strategies employed by children can be sorted into three categories, as defined by Davis et al.: withholding, proactive or no strategies [20].

Proactive strategies are strategies such as faking information, disabling location sharing or having multiple accounts [53]. A 2011 survey in the US showed that 49% of 12-13 year olds admitted to falsifying their age to obtain access to websites and applications [50]. In Davis' research with children aged 10-14 years old, this number was 36%. While falsifying information gives a sense of privacy, bypassing regulations, specifically made for children, is not beneficial: they are actually less protected by bypassing these regulations. A more appropriate proactive strategy was changing profile settings to "friends only", which was mentioned by 64% of children [20]. Disabling location sharing was mentioned by children as well [53]. Similar strategies were found in a qualitative study by Awan et al. [5].

Examples of withholding strategies are to not post certain things or deny friendship requests [20]. Some children indicated that they thought about what they posted and who would see it, as they did not want to post embarrassing things. The 42 children of Davis' study were mostly taught withholding strategies, such as avoiding posting of personal information or talking to strangers. Changing privacy settings was rarely mentioned by their educators (10%) [20].

Overall, children express a feeling of control over their online presence and even experience the internet as a safer space to communicate in than the real world [5]. That they have a sense of control is confirmed by research on co-designing monitoring apps, where 11 out of 26 solutions by children (7-12) were based on self-regulation [59]. However, children do not mind the help of their parents either (15/26 solutions). They did not want their parents to control everything: communication was also important to them. One child thought of buttons so a parent could consult the child before putting up restrictions: "Instead of it being forced and kids having no say, consult [them]". Similar results were found in a co-design paper by Badillo et al [6]. The children in that study introduced several parental control features for the TikTok application, to help them manage stranger danger.

When children were asked by Davis et al. who they seek privacy from, 83% answered "from a known other" (such as a family member). Interestingly enough, children in Mc-

Material nr.	Nr. of ratings	Scenario total	Seen and heard	Only seen	Only heard	Digital
1	116	3	3	0	0	0
2	76	12	6	6	0	0
3	58	5	4	1	0	0
4	55	5	5	0	0	0
5	51	7	4	2	1	0
6	42	1	1	0	0	0
7	31	4	1	0	0	3
8	23	3	2	0	1	0
9	19	1	0	1	0	0
10	14	11	6	2	1	2
Total	485	52	32	12	3	5

Table 2: Stranger danger scenarios analysed from the 10 most popular sources on Teachers pay Teachers [4], categorised by scenarios where the stranger is seen and heard, only seen, only heard or a digital stranger.

Nally's research still allowed parents access to their location (84%), contacts (75%) and browser history (75%). Access to texts and blocking camera access were less agreed upon (42%, 33%) [59]. Children also seek privacy from strangers, which was mentioned by 79% [20]. The next section will go into more detail on the stranger danger phenomenon.

2.2.6 Stranger danger

Stranger danger is a known concept to children and they are taught about it by both parents and educators in detail [4]. It entails the risks surrounding meeting and complying to an unknown person. Historically seen, stranger danger was used to teach children about physical safety [65]. It now extends to the digital world as well. Children themselves indicate to seek privacy from strangers online [20]. It is not surprising that parents' fear of it both on- and offline is highest of all risks: there are news reports that call out predators on TikTok for example, who pose as famous celebrities and ask children as young as 8 years old for nude pictures or videos [94]. Still, it is the risk least likely to occur [14], [40].

It is relevant to discuss what and how children are taught about stranger danger, to understand their knowledge on the topic. On the Teachers pay Teachers (TpT) website, many educational materials on stranger danger can be found that have been created by a community of teachers [4]. For elementary school children, these materials often make use of imaginary scenarios to teach them about not complying to strangers, in order to avoid potential dangerous situations. 10 of the most high-rated lesson materials on stranger danger, rated positively by 485 teachers, were analysed on their example scenarios (52 in total) (see Table 2). The materials were found by searching on "stranger danger" and filtering on elementary school materials.

The researcher of this report made a distinction between scenarios that included a stranger that was seen and heard, a stranger that was seen, a stranger that was only heard and a digital stranger. 32 out of 52 examples included a stranger that was seen and that talked to a child. 12 examples included a stranger that was only seen and 3 only a stranger that was heard. 5 examples were examples of stranger danger in the digital world. One example from each of the ten sources is mentioned in Table 3. There is thus a clear focus on the physical characteristics of a stranger, more so than voice characteristics or digital presence. This could therefore be a reason why children name "stranger danger" less as on online risk [52].

In terms of research, Moran et al. tested children's reactions to strangers using video scenarios. After watching multiple videos, 6 and 8 year olds showed a 55% compliance rate towards the stranger in it. They responded this way mostly out of self-interest but

Material nr.	Example scenario	Category
1	It is raining and a man that you have never met but lives in your neighbourhood asks if you want a ride home from school.	Seen and heard
2	You are walking along the road. A car slows down and starts driving slowly beside you. The driver keeps staring at you.	Only seen
3	A nice-looking stranger approaches you in the park and asks for help finding the stranger's lost dog.	Seen and heard
4	Ben was home alone after school one day, when there was a knock on the door. A nice looking lady was standing outside, asking him to open the door because she had a package for his mom.	Seen and heard
5	Grace and Emily are playing at the park. It is 4:00 p.m. and they were told to be home by 3:30 p.m. A man sitting on the park bench has heard them talking about being late. He offers to give them a ride home.	Seen and heard
6	Someone you don't know waves at you to come and talk to them.	Seen and heard
7	You receive an anonymous text that asks you what your name is and what school you go to.	Digital
8	A stranger approaches you and asks you to come with them to get you a gift.	Seen and heard
9	Someone knocks on the door or rings the bell.	Only heard
10	You are playing a video game online and a player asks for your real name.	Digital

Table 3: One exemplary scenario from every analysed stranger danger lesson [4].

also out of politeness, for reasons as "wanting to be helpful to the adult [in the video]". The 10 year olds in the study showed a 38% compliance rate to the stranger. There were also videos that differed in familiarity of the person and the type of request (request, offer, demand). 10 year olds showed the most nuance in their answers [65]. The research further brings up questions such as when somebody stops being a stranger. It mentioned that knowing a name could already be sufficient.

Co-design research by Badillo et al. revealed that 8 year olds wanted to learn more about stranger danger and withholding techniques. 10-11 year olds wished for more options in managing unwanted contact. Children (N=7, aged 8-11) had to come up with ways to address stranger danger in the TikTok app. 5 out of 7 children were familiar with the application. The children came up with very diverse ideas about preventing oversharing, rejecting strangers and other scenarios [6]. They wished for "Tell a parent" buttons, buttons to warn the police, buttons to "decline" all contact from a person and automatic detection of foul words. Such words would then be replaced with an angry emotion or different text to protect the child from the content. They also wished for fake profiles to be detected and marked by a red dot. Lastly, they wanted an education section within the app to learn more about stranger danger and how to manage this. This research again confirmed that children care about their privacy and want to protect themselves from strangers, something which could be leveraged in the main study of this research.

2.2.7 Giving out personal information

Giving out personal information online is seen as the biggest risk for children [40] (see Table 1). Children themselves are less concerned about this and find it difficult to think about possible consequences of their actions [52], [53]. It is known that children make use of proactive and withholding privacy strategies, yet studies show that they disclose a lot of information. A survey from 2006 found that 82% of teens share their first name on their profile, 79% include pictures of themselves and 61% include their city or town on MySpace [49]. Another research found that 12-14 year olds disclosed more information on MySpace profiles than 15-18 year olds [21]. In general, it has been found that compared to self-disclosure in offline environments, online self-disclosure happens quicker and at deeper levels [8]. According to research, children's self-disclosure is thus high, yet they also show knowledge about privacy strategies and the desire to protect their privacy. The situation can be seen as a privacy paradox, where there seems to be a discrepancy between behaviours and intentions [67].

There are several research examples with children that show this disconnect between the desire to protect their privacy and willingness to share personal information online [20]. For example, children (4-10 years) expressed positive moral judgements about digital tracking and location sharing, even when somebody could track objects that did not belong to him or her [37]. Tracking risks were only mentioned as less as 0.3% of all risks as named by children [52], while this information is privacy sensitive and can be misused. Misuse is closer than may be thought: Snapchat, for example, introduced a location sharing feature so users can always see where their friends are. This can cause even more children to share personal information when peer pressure is considered as well [52].

Miyazaki et al. studied mediating information disclosure by letting children sign up for a website and showing different warnings. They found that an "age-below-13" visual warning and threat of sending an email notification to a parent reduced willingness of children to disclose information online [63]. This finding adds to other research, that reveal the role of parents in children's disclosure decisions. However, Miyazaki's research also showed that disclosure levels in the "visual warning only" condition were even higher than those in a no safeguard condition. This is alarming, since COPPA requires privacy protective measurements such as these to be made. Thus, a website might be in line with COPPA, but still not prevent higher information disclosure [63]. Moreover, personal information is often part of the business model of social networking sites, that specialise in targeted advertisements. For example, the advertising revenue of Facebook in the third quarter of 2020 made up 99% of the company's total revenue from that quarter [43]. While personalised advertisements for children are forbidden by COPPA [1], information about children is still collected and shared with third parties where allowed.

As becomes clear from this section, there is more research that can be done to mediate information disclosure in children. Miyazaki et al. tested out a method for website sign-ups. Whether or not a privacy paradox occurred, was not looked at. The researcher of this report believes it can be beneficial to look at both variables, as they can give insight into children's actions and understanding at the same time. Furthermore, it is important to look beyond the internet as well. So far, the mentioned research focused on applications and websites. However, in today's society, robots are becoming more prevalent. It is relevant to focus on this area of research as well, especially because recent research in this area has raised privacy concerns amongst people. The next section describes research that has tried to map how children make sense of robots and the effects that embodiment can have on information disclosure. There's a gap in the field of Child-Robot Interaction (CRI) when it comes to researching the latter.

2.3 Human Robot Interaction

This section discusses HRI research. It discusses how children view robots, the bonds they create with agents and how embodiment of agents can actually facilitate more information disclosure. Not much research has been done on information disclosure and embodiment of agents with children. Therefore, insight is gained from papers that are available, including adult research.

As mentioned before, robots are increasingly found around us: in schools, stores, hotels and more. The field of Human Robot Interaction is relevant to discuss, because our privacy is not only susceptible to our mobile devices, but also to robots. A robot is simply another form that can collect personal information from its users.

Robots are defined as programmable machines, capable of carrying out a sequence of actions automatically [22]. Their control is embedded within the machine itself or they can be controlled from the outside. There are many different kinds of robots, but one type of robot that is focused on in particular in this report, is the social robot. Duffy et al. defined a social robot as a physical entity embodied in a social environment, that is sufficiently empowered to achieve its own goals and those of its community [24]. When a social robot can speak and has a physical form, it can also be regarded as an embodied conversational agent (ECA). Its physical form must be capable of allowing non-verbal communication, too [16].

In this report, smart speakers such as Amazon's Alexa and Google's Google Home are regarded as social robots that are ECAs, too. Although their embodiment is simple, it includes sensors that indicate when the device is listening and "thinking" (loading) for example. They are also capable of non-verbal communication, which means communication outside of words. For example, loudness or tone of voice, both paralinguistic cues, can be controlled.

Even though ECAs have been used in the experiment of this report, it must be noted that not every robot mentioned in the upcoming section is an ECA. Furthermore, not every conversational agent that is mentioned is embodied and may be regarded as a social robot.

The next section pays attention to how children react to different types of robots and agents, as this creates insight into how children might react to the conversational agents from the main study in this report. The effects of embodiment of agents on information disclosure are discussed in the section after that. However, those effects are mostly reported using research on adults: there's a gap in the field of CRI in regards to this topic.

2.3.1 How children view robots

It has been found through research that children view robots as different "beings" than computers [44]. As opposed to computers, children attribute more intent and emotion to robots, especially those they can interact with socially and psychologically [80]. This attribution starts from the age of 3 to 5 years old, wherein children develop a sense to perceive emotional and mental states of others. They apply this "Theory of Mind" to make sense of the world, whether it be about the state of their pet, the moods of their parents or robots [89]. Even before this development, children might have been in contact with a robot already. Participants of Sciuto et al.'s in-home interviews in 2018 mentioned that interaction between child and robot happened from as young as 15 months. In particular, the parents of the 15-month old son told them that the child knew where to look to find Alexa when a family member interacted with it. Another child, a 2.5 year old, assumed that Alexa could see the colours of her crayons when she asked Alexa to name them, even though Alexa has no facial features. Most parents mentioned that their children love asking questions to the embodied conversational agent [74].

That children can create positive bonds with different types of agents is seen in a study by Druga et al. [23]. They let children (N = 26) interact with Alexa, Google Home, Cozmo (an autonomous, small robot toy) and the Julie Chatbot (a conversational agent - chatbot - on the internet). 70% of the 3-10 year olds described every agent as "friendly". 60% or more was attributed to "truthfulness" for every agent. When it came to intelligence, the younger children (3-5) believed they were smarter than most devices or were neutral about it. Older children believed the device was smarter, with the biggest difference in opinion on Alexa (100% in the older group believed this, versus 20% in the younger group). Differences in agents as well as differences in the age of a child itself thus play a role in these beliefs.

According to Williams et al., education also plays a role in how children perceive robots [91]. In their study, children were taught about AI. They found that children who performed worse on AI assessments after lessons, believed robots to be less smart than themselves, while children who performed well saw robots as "people who were smarter than them".

Children can form strong bonds with robots, even when they resemble a living organism only slightly. In a research with the robot dog AIBO, 60% of 72 children between ages 7-15 affirmed that AIBO had mental states, sociability and moral standing [60]. Moral standing means that the way people treat something or somebody makes a moral difference. It has been shown that a mistreated robot can cause distress in children, such as when a robot is put in the closet by a researcher, even when it claims it is "scared of the closet" [45]. Children found it morally wrong to force the robot against its will, but at the same time indicated they would not grant a robot civil rights or entitlement to its own liberty. It seems that robots are not held to the same standard as humans, but come pretty close.

The fact that children form such strong bonds with robots makes them vulnerable as well. Research done by Vollmer et al. highlights this: they found that children are more likely to conform to a group of three robots that give the wrong answers to a task than adults are [85]. The children in the study were between 7-9 years old. The group of social robots were thus able to apply social pressure to a child, having it make the wrong decision. The trust that is given to robots is thus not always a good thing. In the next section, several papers on the embodiment of robots and information disclosure are discussed, as these are topics that are integral to this report.

2.3.2 Embodiment and information disclosure

Robots are a separate category of research that challenge traditional research findings. They pose challenges to user privacy because of their embodiment and novelty for users [10]. Current privacy legislation does not take the possible impact of a physical appearance into account: recent research suggests that it should, because the embodiment of a robot can change users' privacy considerations [84]. "Embodied" usually refers to body parts, bodily actions, or body representations [36]. In the case of robots/agents, this often means that they have human-like features, such as a head, arms, etcetera. Usually, embodied entities are dynamic instead of static, so they can interact with the real-world environment through sensors and motors.

There are several papers worth discussing that looked into the effects of embodiment. Many of them do not have child participants, but are still worthwhile to mention. First, a paper by Vitale et al., looked at how embodied agents had an effect on people's willingness to disclose private information. It turns out that people were more willing to give information to a robot (which carried a tablet for interaction purposes), than only a tablet [84]. However, this was only the case in the transparent conditions, where both systems were clear about privacy guidelines. When the guidelines were not presented, there was no significant difference in sharing information. The researchers thus concluded that the physical design of the system had an impact on users' decisions.

Another privacy related study comes from Caine et al. They studied how robots with different degrees of embodiment would influence privacy enhancing behaviours of older adults [15]. There was one condition with a simple mounted camera, one with an immobile "embodied" robot with a camera and a mobile embodied robot with a camera. The adults had to prepare a secret birthday surprise for their caretaker, involving changing of shirts, hanging up decorations and inviting people through the phone. They got told that their caretaker might be looking at the camera footage. The researchers found that the participants showed more behaviours to enhance their privacy in the simple webcam condition than with the embodied robots. The exact reason for this finding could not be completely explained by the researchers.

Besides physical presence, physical touch can facilitate disclosure too. Touch is important in creating meaningful human connections. Shiomi et al. used this principle to see whether participants would disclose more personal information to a robot after hugging it, as opposed to not hugging it [75]. The participants were instructed to talk about their own life to the robot. They disclosed more personal information and talked to the robot for a longer amount of time, indicating a stronger bond.

Whereas Shiomi's robot did not disclose information about itself, research by Moon showed that if an agent disclosed information about itself as well, it helped in getting people to reveal more about themselves, too [64]. This phenomenon is called reciprocal self-disclosure. Besides this, they found that participants were more likely to disclose intimate information when they were first "warmed up" with some introductory questions. The sequence of questions can thus make a difference in disclosure, too.

Embodied agents are increasingly anthropomorphic, which means that human characteristics are attributed to it. The previous section contained many examples of how children anthropomorphise agents. But the fact that they are not actually humans, also influences our decisions. For example, Lucas et al. found that patients had increased willingness to disclose personal information to autonomous agents than to tele-operated ones [54]. If they were certain that the virtual human on the screen was an agent instead of a real human behind the scenes, they were less anxious about being evaluated and making a bad impression. They would show more sadness to the autonomous agent as well. Agents may profit from this "anthropomorphic advantage". Sannon et al., for example, used a 3x3 experiment differing in degrees of sociability of chatbots on the internet and types of data sharing practices. Participants felt more negatively towards agents who would share their data outside of the company with third parties [73]. However, they were less negative about this when they were interacting with the most social chatbot.

Embodied conversational agent Alexa also compromises privacy of users in a way. Major et al. found that users often do not understand that some skills are run by third parties and when these are in effect. Surprisingly, experienced users were even more likely to mistake third party skills for native Amazon functionality. One skill that stood out in being recognised correctly as a third party skill, was Jeopardy (79.7% of participants recognised this). The researchers believe that the change in voice from Alexa to Alex Trebek facilitated this recognition and suggest future research to look into this [56]. In conclusion, Alexa users have bad conceptual models about where and with whom their information ends up if no cues or additional information are given. Don Norman defines the conceptual model as a person's mental model of how a product works: "The design should project all the information needed to create a good conceptual model of the system, leading to an understanding and feeling of control" [68].

That robots can get away with more, even in privacy sensitive settings has not been researched much in relation to children. However, Leite et al. conducted research in which a robot would reveal a secret that only the child knew about. Children in this condition noticed the revelation and showed emotional responses to it, with younger children (4 to 6 years old) being more affected than older children (7-10 years old) [48]. However, even though negative affect was found in the revelation moment and descriptions of parents about their children's responses (unsure, confused, shocked, surprised), the moment did not have an impact on the likeability of the robot and willingness to interact with it again.

Most of the mentioned research was done with adults and showed an increase in information disclosure with embodied agents. That rules and regulations do not take this seeming advantage into account, is a legitimate concern. Furthermore, it becomes even more relevant to research information disclosure and embodied agents with children: children may be extra vulnerable to this, as the previous section showed that they create very strong bonds with robots.

2.4 Literature summary

This section summarises all the literature that is discussed above. It identifies the gaps in knowledge and links the discussed literature to craft a novel research question.

Firstly, the literature study has focused on privacy matters such as policies and regulations. It is well known that most policies are lengthy and full of jargon [3], [35], [46], [57], [61], [70], [76]. Attempts have been made to change the presentation of these documents, with some of them actually improving users' privacy awareness [19], [26], [46], [77]. Promising research made use of timing to show smaller privacy policy parts [26]. It has become common for applications and websites to use something similar, named privacy permission requests. Such requests pop up when certain information becomes relevant to collect and they contain explanation on what happens with the collected information. They could be suitable for raising privacy awareness in children, as they are compact.

Furthermore, the literature study looked at how children use technology and view privacy matters. Most privacy regulations protect children until the age of 13. However, children go online unsupervised from 8 years old [12]. Therefore, the age group of 8-12 year olds find themselves in a grey area, making it relevant to study them in more detail. Previous research showed that children develop their own protective strategies and are generally privacy aware [5], [20], [50], [53]. Most children are taught by their educators and parents about risks online. Still, they give out a lot of personal information, which is also the biggest risk online for children [40]. The situation can be seen as a privacy paradox [20], [67]. Some researchers have tried to lower information disclosure. Miyazaki et al. found significant results for this through a sign-up experiment that gave parental threats and age warnings to children [63]. However, different approaches might be interesting to explore as well.

It was further identified that parents fear "stranger danger" and how children are made aware of this risk from a young age. In order to minimise this risk, children are taught not to comply to requests from unknown persons [4], [65]. Stranger danger now also extends to the digital world and children have shared that they seek privacy from strangers online [20]. They recognise that they can encounter stranger danger online, but do not necessarily take potential dangers, caused by strangers behind a company, into account. In general, they are very trusting of companies up until the age of 12 years old [53]. Using children's knowledge on "stranger danger" could offer an interesting approach to researching information disclosure and privacy awareness.

Lastly, HRI research was looked at. Children can create strong bonds with robots and almost regard them as humans [44], [45], [60], [74], which can be leveraged positively but also makes them vulnerable [48], [85]. Furthermore, the degree of embodiment can have effects on people's behaviours [15], [54], [86], causing even higher information disclosure with higher degrees of embodiment [84]. These findings raise concerns about robots and privacy, as no regulations currently take into account the seeming advantage of embodiment. Also, most studies on information disclosure and embodiment have been done with adults, which means there is a research gap that can be filled. It can be argued that research with children on this topic is very necessary, precisely because they are capable of creating such strong bonds with robots. Furthermore, the embodiment of robots/agents affords a fusion with the phenomenon "stranger danger", to turn an agent into a stranger.

In conclusion, a novel research experiment can be set up where children have a conversation with an agent. There can be multiple agents with different levels of embodiment to study their effects on information disclosure and privacy awareness of children. Furthermore, as agents can facilitate a change into a "stranger presence", the effects of this new approach can be measured for information disclosure and privacy awareness as well. Such a change might happen at moments that privacy sensitive information is asked, as the metaphor behind the stranger can be a company whose "stranger employees" collect information. The asking of privacy sensitive information can be done using privacy permission requests. Children's privacy awareness can be measured by their understanding of these requests and their information disclosure can be measured by their compliance to the requests. Measuring both these variables can give more insight into the privacy paradox that children are susceptible to.

The next section will introduce the research question that followed from this literature study.

3 Research Question

The following research question is proposed:

> How does the level of embodiment of an embodied conversational agent and the (non)presence of a stranger within this agent during privacy permission requests, influence information disclosure and privacy awareness of children?

The research question specifically uses the term embodied conversational agent (ECA). The term ECA offers clarity on the experiment itself too, as children will *converse* with an *embodied agent.*

The research question differentiates between a "high" and a "low" level of embodiment. In this study, "embodiment" encapsulates having a physical, tangible "body". High embodiment is defined as having more anthropomorphic - human-like - features. Low embodiment in this study entails that the agent is not very anthropomorphic: it does not have clear human-like features. However, it can use and understand human speech.

The presence or absence of a stranger within the agent is defined as follows: a stranger presence within the agent means that the agent "embodies" a stranger during privacy permission requests in the conversation. A stranger is defined as a person that is unfamiliar or that does not belong in a particular context [65]. Classifying a person as unfamiliar is often done based on whether their physical characteristics or voice characteristics are recognised or not [4]. In the context of this research, a stranger presence is facilitated by a change in these characteristics of the agent. This can either mean an auditory change (voice), or both an auditory and a visual change (voice and face), depending on the affordances of the ECA. Next to this, knowing a strangers' name can already ensure that they are not seen as a stranger anymore [65]. The stranger presence thus does not share any personal details about themselves in the conversation with the child. Lastly, the "no stranger presence" means that the agent stays the same during the entire conversation with the child. This agent does share personal information throughout the conversation, although not during the privacy permission requests either.

4 Hypotheses

The research question is split up into multiple hypotheses, that are as follows:

4.1 On information disclosure (H1)

Hypothesis H1A: Children will disclose more information to a "no stranger presence" compared to a "stranger presence" within the ECA.

A child can create a strong bond with a conversational agent [23], [60], [74]. It is believed that they will create a stronger bond with a conversational agent when a stranger presence within the agent does not occur, because more time spent with "one" agent appearance can strengthen the bond [44], making the child more vulnerable to conform to the privacy permission requests [85].

Besides this, it is known that an unexpected event can cause negative affect in children. In Leite's research, children were, according to their parents, unsure, confused, shocked and surprised that a robot knew a secret that it was not supposed to know [48]. In this research, the unexpected event is a stranger presence: children might be surprised negatively by this presence and feel less willing to share information.

The stranger also does not share any personal information about itself, which is suspected to contribute to children feeling less willing to disclose information. When an agent discloses information about itself as well, it can help in getting people to reveal more about themselves too [64]. As the “no stranger” does share personal information, information disclosure might be higher in “no stranger presence” conditions.

Lastly, children are taught not to comply to strangers [65]. In Miyazaki's research, children disclosed less information when a parental threat was given [63]. In this case, a stranger “threat” occurs. From Davis' research, it is known that children seek almost as much privacy from strangers as from their parents [20]. This can be another argument to assume that children will disclose less information to a stranger presence within an agent.

These arguments lead the researcher to believe that children will disclose less information to a stranger presence within the agent, compared to a “no stranger presence”.

Hypothesis H1B: Children will disclose more information to a “no stranger presence” within a high embodied agent, but less to a “stranger presence” within a high embodied agent, compared to a low embodied agent. An interaction effect thus occurs between the level of embodiment of an agent and the stranger (non)presence within an agent.

It is predicted that a high embodied agent will lead to a higher disclosure of information than a low embodied agent, when no stranger presence occurs. This is based upon findings that state that people share more with agents that are more social and higher embodied [73], [84]. However, it is expected that children will disclose less when a high embodied agent turns into a stranger compared to when a low embodied agent does. A stranger is usually recognised by their unfamiliar characteristics. These characteristics include physical characteristics and voice characteristics. This can be seen in research approaches [65], but also in educational materials [4]. Especially educational material on stranger danger focuses on scenarios where the stranger is seen as well as heard. The multi-modal experience of the high embodied agent might therefore match children's expectations of a stranger better and have a bigger impact on them. This would then, in turn, lead to less information disclosure (see Figure 1).

4.2 On privacy awareness (H2)

Hypothesis H2A: Children will show higher privacy awareness when an ECA has a high level of embodiment, regardless of a stranger (non)presence within the agent.

Throughout their youth, children are taught to listen to other humans, such as their teachers, parents and other adults [65]. Since the high embodied agent has more human-like features, it is thought that children will listen more carefully to this agent. Another reason for more attention could be that the high embodied agent can be viewed as more visually interesting and dynamic. Its face and eyes move while talking and it also has a camera that tracks the child, so it follows the child if they move. These are a things that could maintain the attention of a child.

The low embodied agent is quite static: it only produces sound and shows a few inner states (such as listening and loading) using white LEDs. It is thus not very stimulating.

Following these arguments, it is suspected that children will pay more attention to the permission requests of high embodied agents, leading to higher privacy awareness.

Hypothesis H2B: Children will show higher privacy awareness when an ECA has a “stranger presence”, regardless of the level of embodiment of the agent.

Children are taught about stranger danger from a young age [65]. Because it is not always clear what the intentions of a stranger are, they are taught to refrain from contact with them and not comply to their questions [4], [14], [20]. It is thought that a stranger presence within an agent will attract more attention from children, because they will experience an agent presence that they are unfamiliar with and that they need to deal with. It is assumed that this attention will be directed to the permission request as well, in order to make a decision that minimises stranger danger. Their attention towards the privacy permission request would then lead to higher privacy awareness in the stranger presence conditions. See Figure 2 for a visual explanation of hypothesis 2.

4.3 Information disclosure and privacy awareness correlation (H3)

Hypothesis H3: There is a negative correlation between privacy awareness and information disclosure.

It is assumed that children are relatively privacy aware, based on the literature study. Despite their privacy awareness, information disclosure still happens. A negative correlation between these two variables is hypothesised, which can be explained by the hypothesised differences between the no stranger presence and stranger presence conditions. The negative correlation means that a higher level of privacy awareness leads to lower information disclosure.

The stranger conditions are hypothesised to show low information disclosure and high privacy awareness. Several arguments are used to hypothesise the low information disclosure, such as the negative emotions connected to the unexpected event, a less tight bond with the agent due to its stranger presence and the activated knowledge on not complying to strangers' requests, to avoid stranger danger. Furthermore, the high privacy awareness score is hypothesised to occur because a stranger presence might facilitate more attention from children. This could make them listen and remember the privacy permission requests better, in order to minimise stranger danger. The no stranger conditions, in comparison, are hypothesised to show higher information disclosure and lower privacy awareness. The differences between these condition groups would then lead to a negative correlation between privacy awareness and information disclosure. See Figure 3 for a visual explanation of hypothesis 3.

4.4 Summarised hypotheses

In summation, these are all the hypotheses:

H1A Children will disclose more information to a "no stranger presence" compared to a "stranger presence" within the ECA.

H1B Children will disclose more information to a "no stranger presence" within a high embodied agent, but less to a "stranger presence" within a high embodied agent, compared to a low embodied agent. An interaction effect thus occurs between the level of embodiment of an agent and the stranger (non)presence within an agent.

H2A Children will show higher privacy awareness when an ECA has a high level of embodiment, regardless of a stranger (non)presence within the agent.

H2B Children will show higher privacy awareness when an ECA has a "stranger presence", regardless of the level of embodiment of the agent.

H3 There is a negative correlation between privacy awareness and information disclosure.

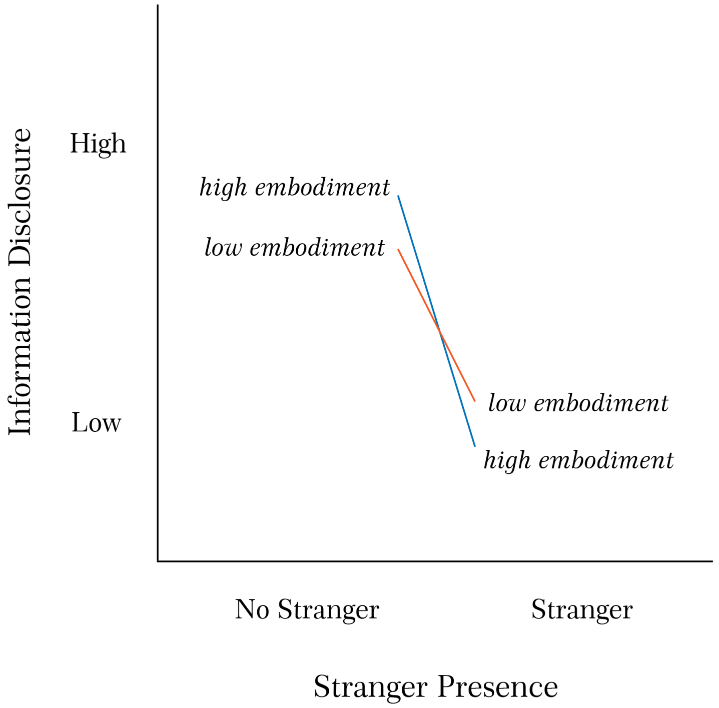


Figure 1: Hypothesis 1 on information disclosure.

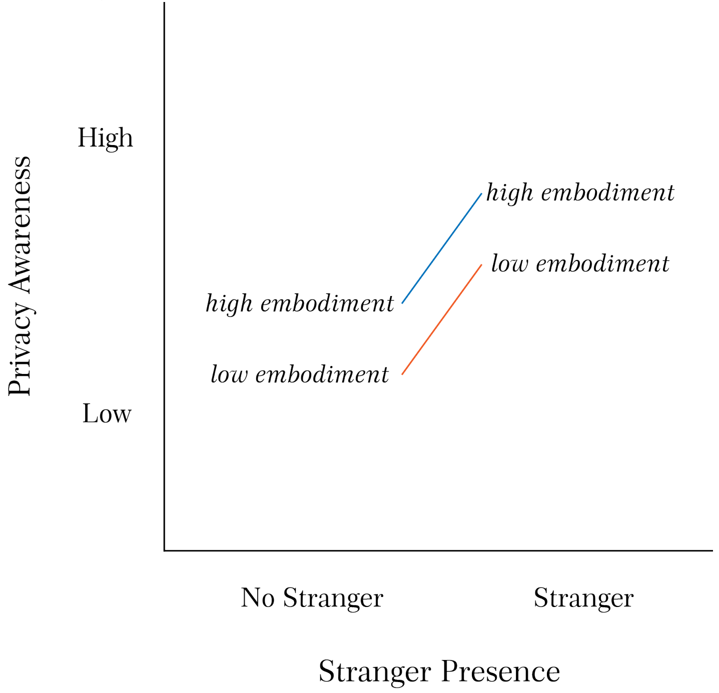


Figure 2: Hypothesis 2 on privacy awareness.

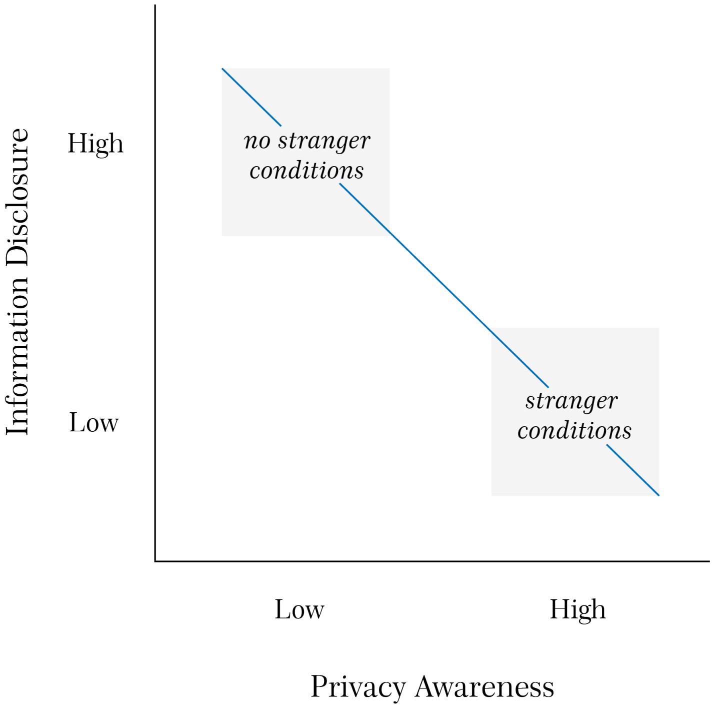


Figure 3: Hypothesis 3 on the correlation between privacy awareness and information disclosure.

4.5 Approach

In order to answer the research questions, multiple pilot studies were set up before starting with the main study. Through the pilot studies, results were obtained for the design of privacy permission requests. Furthermore, the recognition of auditory and visual changes was tested to see whether children could pass part of the manipulation check. Through the main study, data was obtained to answer the research questions.

5 First pilot study: privacy sensitive variables

5.1 Goal

The first pilot study provided insight into children's views on privacy and their experiences with robots. Most importantly, multiple choice questions identified the perceived privacy sensitivity of sixteen variables, which helped determine the variables requested by the ECAs in the main study. Not much research has been done that asked children about the privacy sensitivity of certain variables when giving these to an application or website. Past research asked adults' views on giving out these variables of children [59] and discussed similar privacy matters with children through interviewing, obtaining qualitative instead of quantitative data [20], [49], [52]. The pilot study got approved by the ethical board.

5.2 Participants

12 children (M = 3, F = 8, O = 1) aged 8-12 years old joined the experiment (8 = 25%, 9 = 8.3%, 10 = 16.7%, 11 = 41.7%, 12 = 8.3%). The distribution of participants is thus skewed towards eleven year olds and girls as well. The participants were either found through parents who are acquaintances of the researcher or through the parent portal of a daycare at the University of Twente.

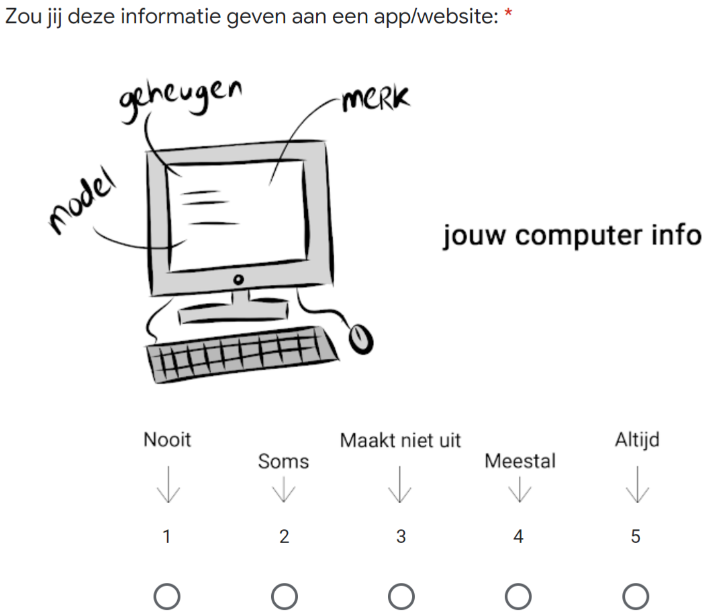
5.3 Method

The research took around ten to fifteen minutes and was filled in at home, because of COVID-19 safety measures. The thesis website contains all materials used in the pilot study [97].

The survey contained explanation videos of the researcher as well as drawn images to make the survey fun and easily understandable for the children who participated (see Figure 4 for an example question). In the videos, the researcher made it clear that the child was the expert, that there were no right or wrong answers and that they should not let themselves be influenced by their parents, if they were present.

The first part of the survey included background questions such as what devices children own, what applications they visit most, whether they have ever read a privacy policy and how familiar they are with robots. The second part of the questionnaire revolved around identifying what children find privacy sensitive information. Ackerman et al. researched something similar [2]. They asked adults what concrete information they would give away to websites, such as age, name and income. They were also asked to fill this in from the viewpoint of having a child between 8-12 years. Furthermore, in the co-design research of McNally et al., children were asked what features they would allow their parents access to, such as picture access, camera access and access to contacts [59]. This pilot study combined the variables of these studies and asked these to children themselves. The children had to imagine whether or not they would give this information to a website or application. The variables that were asked about can be found in Table 4.

Variables that are stricken-through are discarded variables: decisions on leaving out certain variables from the survey were made depending on whether children would be aware of a variable (thus discarding social security number), were owners of a variable (discarding credit card info) and whether the information was usually obtained by websites and applications or not (removing restrict internet access). Some variables were rephrased to be more relevant to children, such as changing income to pocket money. Redundant variables were discarded as well (such as "view search history", keeping "view browsing history"). One variable was added, namely "picture access", as it is commonly asked by applications and websites. In the end, sixteen variables were presented with a five-point Likert scale ("would never give out" to "would always give out") in the survey for children (see Table 4).



Zou jij deze informatie geven aan een app/website: *

Nooit	Soms	Maakt niet uit	Meestal	Altijd
1	2	3	4	5
○	○	○	○	○

Figure 4: Example of a survey question about a privacy sensitive variable. *Translation: Would you give the following information to an app/website: your computer info?* The questions used a five-point Likert scale (from never to always).

From Ackerman's research	From McNally's research
~~Social Security Number~~	~~Block~~ Camera Access
~~Credit Card Info~~	View Text Messages
~~Income~~ Pocket Money	~~Restrict Internet Access~~
Phone Number	~~View Social Media Posts~~
Medical Info	View Browsing History
Address	~~Block App Downloads~~
Full name	~~View Call Logs~~
Age	~~View Search History~~
Email Address	View Contacts
Computer Info	GPS Location Tracking
Favourite Snack	
Favourite TV Show	

Table 4: Variables as chosen from existing research that were asked of the children in the pilot study. The variables are ranked from least to most accepted, respective to each study [2], [59].

5.4 Results

According to the results, the children in the pilot study sample were experienced with technology: only one out of 12 children did not own a device by themselves. The other

11 children owned a device, with 6 of them owning more than one device. 91.7% usually went online without supervision and 83.3% of all children had lied before when signing up for a website.

83.3% of all children indicated that they found privacy important. Several reasons included: “Yes, because it is personal”; “Yes, because they are not allowed to know everything.”, “Yes, because they are things that are yours and not someone else’s and otherwise people will steal your things”, “Yes, for the wrong people.” and “Otherwise everyone can see that it is you.” Answers on this question were often a bit ambiguous in the meaning of “it” and “they”. Presumably, “they” indicates people with bad intentions, possibly strangers. Further explanation was not asked for.

Answers were more divided on whether the children knew about privacy policies and had read them before. 25% had read a privacy policy before, 33.3% had heard of privacy policies, but never read them and 41.7% had never heard of them before.

When it came to robots, 33.3% had never interacted with a robot before. The other 66.7% had experience with robots, mostly with remote-controlled toys. 25% of all children had experience in talking to an embodied conversational agent. None of the children lived in a household that owned a smart-speaker such as a Google Home or an Alexa.

Owning multiple devices did not mean that the children had more often heard about or read privacy policies. They also were not more experienced with robots.

The results on giving out personal information are shown in Table 5. Children were the least protective of their email address and favourite snack. They were very unwilling to give out access to their contacts and medical info. An interesting finding is that each variable received low information disclosure scores (a “would never disclose” or “would sometimes disclose” score) from at least half of all participants. The children of this sample thus score quite low on information disclosure.

Privacy variable	Mean value (SD)
Email Address	2.92 (1.165)
Favourite Snack	2.75 (2.750)
Full Name	2.67 (1.231)
Age	2.58 (1.379)
Favourite TV Show	2.42 (1.564)
Picture Access	2.00 (1.128)
Camera Access	1.92 (0.900)
Address	1.75 (0.866)
Phone Number	1.75 (0.965)
Pocket Money	1.58 (0.900)
GPS Location Tracking	1.58 (0.669)
Computer Info	1.42 (0.793)
View Browsing History	1.25 (0.452)
View Text Messages	1.25 (0.622)
View Contacts	1.17 (0.389)
Medical Info	1.08 (0.289)

Table 5: Information disclosure mean values (standard deviations) for each privacy sensitive variable (1: would never disclose, 5: would always disclose), ranked from high to low disclosure scores.

5.5 Discussion

The ranking as found through this pilot (see Table 5) is used in deciding upon the variables of the permission requests in the main study. The found averages were divided into three categories, determining the least to most privacy sensitive categories. Two variables were picked from the most privacy sensitive category, the other categories provided one variable each.

Only after conducting the first pilot study, the two ECAs that would be used in the main study were known: a Furhat robot and a Google Home Mini (see Section 7.4 for further explanation). Therefore, the researcher also looked whether or not a variable would provide a realistic scenario in a conversation between these ECAs and a child. For example, picture access would be difficult to insert into a conversation: how could actual access to this be given? An alternative to full access would be to have a child show a few pictures to the agent. This would create several obstacles: 1) not all children own mobile phones, 2) the Google Home Mini, having no camera system, would not be able to recognise the content and 3) the Furhat robot would require advanced computer vision to recognise the content in order to make an appropriate comment on it. Following these lines of reasoning, variables were either chosen or discarded. In conclusion, the variables decided upon for the main study were name (least privacy sensitive), phone number (moderately privacy sensitive), computer info and access to contacts (both identified as very privacy sensitive). The creation of the privacy permission requests can be found in Section 7.6 on conversation design.

5.6 Limitations and future research

In regards to limitations, a limitation to be mentioned is that the questions are purely hypothetical. It is therefore unknown whether children's behaviour would match their intentions in a real scenario. There is often a difference between these two in the case of information disclosure [67]. Although parents were instructed to not to influence the child, they might have or their presence might have.

Furthermore, the children were asked what their answers would be in the case of a website or an application asking for information. Their answers could have differed in the case of an ECA. Since it was not completely certain which ECAs were to be used at the time of starting the pilot study, websites/applications seemed a sufficient second option.

Lastly, a survey format has limitations in regards to exploring answers to open-ended questions in a deeper way, since the answer is only known after submitting the survey. A video interview format might have worked better to explore children's views on privacy further.

6 Second pilot study: auditory and visual changes

6.1 Goal

This small pilot study identified whether auditory and visual changes within the ECAs were recognised by children, before starting the main study. The recognition of these changes was important, since these changes are part of creating the “stranger presence”.

6.2 Participants

4 children (M = 2, F = 2) aged 9-12 years old joined the study. Every age in the sample was represented by one child. The children were from the same group as the first pilot study. This study has no relation to the first pilot study. Furthermore, the children did not know anything about the main study either. An unbiased opinion was therefore still possible.

6.3 Method

The children were sent four files. Two files were separate pictures of the agents that were used in the main study. The other two files were an audio file of the Google Home Mini and a video file of the virtual Furhat robot interface. The sent fragments can be found on the thesis website [97]. In each fragment, the agent told a short part from a fairy tale. In the middle of each fragment, a change would occur: depending on the agent, an auditory or auditory and visual change took place. The children were not told about the change, since the objective was to find out whether the change would be noticed by them or not.

The children were told to watch and listen to the fragments and comment on everything that they observed. Their parents made a small video or wrote in text what their children said about the fragments, which they sent to the researcher.

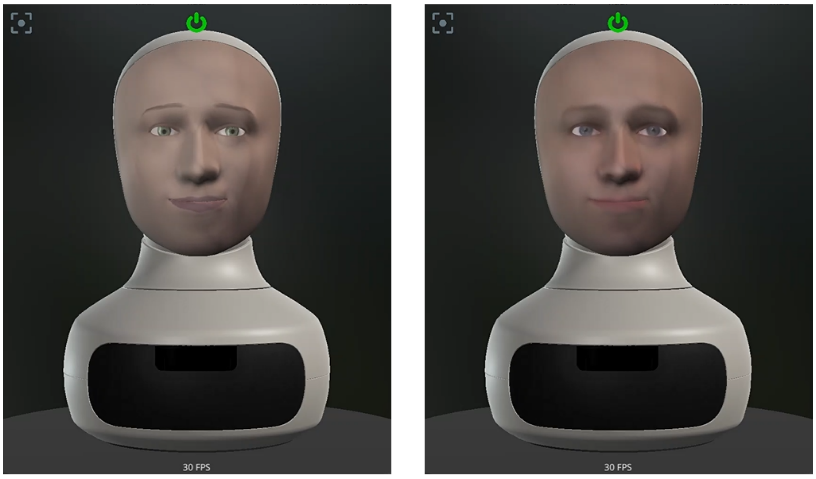
6.4 Motivation behind changes

The main research question differentiates between a “stranger presence” and a “no stranger presence”. One of the objectives of the research is thus to find out whether a stranger presence during privacy permission requests can facilitate lower information disclosure scores and higher privacy awareness.

As mentioned before, a stranger is often described as an unknown person, that does not need to stand out to be a stranger [4]. They do not have any “rare” features, such as blue skin, a missing eye or a bloody nose, for example. The “stranger presence” within the agent should therefore be kept quite “normal” as well. It should not push the child towards certain answers because of its looks.

The changes that were tried out in this pilot study were thus realistic, instead of exaggerated. It was decided first that both ECAs should be male agents: the male voices between the two programming platforms were easier to match in tone than the female voices. As a result, the Furhat robot also has a male face to match its voice.

The auditory change of the stranger presence was a lower voice. As for the visual change, the Furhat robot changed to another male face (see Figure 5 for this change). The researcher thought that the “stranger” face of the Furhat looked a bit older than the “no stranger” face, therefore it was matched with the lower voice.



No Stranger / Rob

Stranger / Robot B.V.

Figure 5: The visual change of the Furhat robot.

6.5 Results

6.5.1 Furhat video

Participant 1 (F, 10) was impressed by the Furhat and asked “Wow, can he actually say all of this?”. She noticed the face as well as the voice change in the robot. Participant 2 (M, 12) noticed the face change and also the voice change, although he was more certain about the face change. He enjoyed the voice of the Furhat more than that of the Google Home Mini, finding the Furhat less monotone. Participant 3 (M, 9) and 4 (F, 11) were brother and sister and looked at the fragments at the same time. They noticed that the face changed and found that the Furhat talked very fast.

6.5.2 Google Home Mini audio

Participant 1 noticed that the middle part changed in voice but she also had the feeling that the fairy tale was missing parts, which was not the case. Participant 2 found the voice monotone. He noticed that the fragment was split up into three parts and found that all the parts sounded kind of different. Participant 3 and 4 did not notice any change in voice. They said that it felt as if the voice took no breaks between sentences.

6.6 Discussion

The results gave insight into what changes needed to be made to create clear auditory and visual changes in the agents. Firstly, the pilot study showed that the voice of the Furhat and Google Home Mini are perceived differently, with the Furhat voice being perceived as more pleasant. Due to the different frameworks of the agents, it is not possible for both agents to have exactly the same voice: the researcher tried to match them as well as possible. What mostly caused the different experience between the two agents in voice, was the pauses between sentences. These pauses had to become clearer for the Google Home Mini, to avoid being “monotone”.

6.7 Limitations and future research

Due to COVID-19 and working from home, the researcher was unable to create a video with the actual Furhat robot. This is why separate pictures of the agents were sent, to give an idea of the physical appearance. Of course, the experience would have been different if the children could have experienced the agents in real life. However, for the objective of the small study, this approach gave enough insight and was deemed sufficient.

As mentioned in the discussion, the pauses between agent sentences had to be made clearer for the main study. Therefore, each sentence of text of the Google Home Mini in the main study, was wrapped in a SSML sentence command, that made the start and end of sentences more clear. The auditory change in the Google Home Mini was also made more distinctive for the main study. The voice of the Furhat was therefore tweaked as well to not introduce too big of a difference between the embodiment conditions. For both agents, the voices were slowed down too for a better experience in the main study.

Lastly, in the main study, not only the changes within the agent need to be noticed, but the resulting presence should be recognised as a stranger. Sections 7.6 and 7.8.4 on conversation design and the manipulation check address this further.

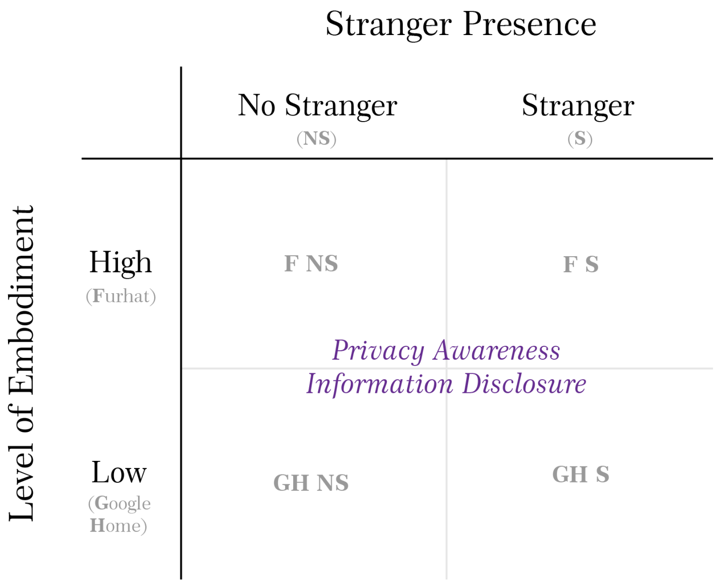
7 Main study

7.1 Goal

The goal of the main study was to provide answers to the research question, namely how the level of embodiment of an embodied conversational agent and a stranger (non)presence within this agent during privacy permission requests influences the information disclosure and privacy awareness of children.

7.2 Study design

The study is a 2x2 factorial between-subject design. The first independent variable is the level of embodiment and is represented by the Furhat (high embodiment) and Google Home Mini (low embodiment). The second independent variable is a stranger presence or a no stranger presence. This leads to four conditions: a visualisation of the experiment design can be found in Figure 6. The following sections will explain the choices of the agents and the extent of the stranger presence in more detail.



		Stranger Presence	
		No Stranger (NS)	Stranger (S)
Level of Embodiment	High (Furhat)	F NS	F S
	Low (Google Home)	GH NS	GH S

Figure 6: Set up of the 2x2 factorial design with independent variables as table headers. The cell contents display the four conditions: Furhat No Stranger (F NS), Furhat Stranger (F S), Google Home No Stranger (GH NS) and Google Home Stranger (GH S). The dependent variables are written in purple text.

7.3 Participants

Children from two primary schools in Enschede participated in the research. 86 children were given consent by their parents and 83 of those children gave consent as well. 79 children completed the full experiment (F NS = 19, F S = 21, GH NS = 19, GH S = 20). Some entries had to be removed due to technical difficulties (failing WiFi or agent sensors, for example). Another reason for removal was some children's lack of motivation to finish the questionnaire. Next, the entries that did not pass the manipulation check (see Section 7.6.5) were removed, after which 60 participants were left in the data pool. The distribution of age, gender and conditions of the 60 participants can be found in

Table 6. 80% of the children in the sample had never spoken with a robot before, 16.7% had spoken with a robot before and 3.3% did not know if they had.

Condition	N	Gender	N	Age	N
F NS	19	M	26	8	18
F S	15	F	33	9	15
GH NS	20	O	1	10	18
GH S	6			11	8
				12	1

Table 6: Condition, gender and age distribution in the main study (N=60).

7.4 Materials

Two ECAs were used in the experiment to reflect the different levels of embodiment (see Figure 7). The Google Home Mini is a smart speaker, that represented the low embodied agent. It can use and understand human speech, but does not have any other human-like features. It can show a few internal states, such as listening, through its three LEDs on top of the speaker.

The Furhat robot is a bust with a projected face on it, that represented the high embodied agent. It can use and understand human speech. It also has a human-like face: its eyes and mouth move during talking and it can keep eye contact with the user as well. The two agents were programmed to operate autonomously during the conversation with the child.

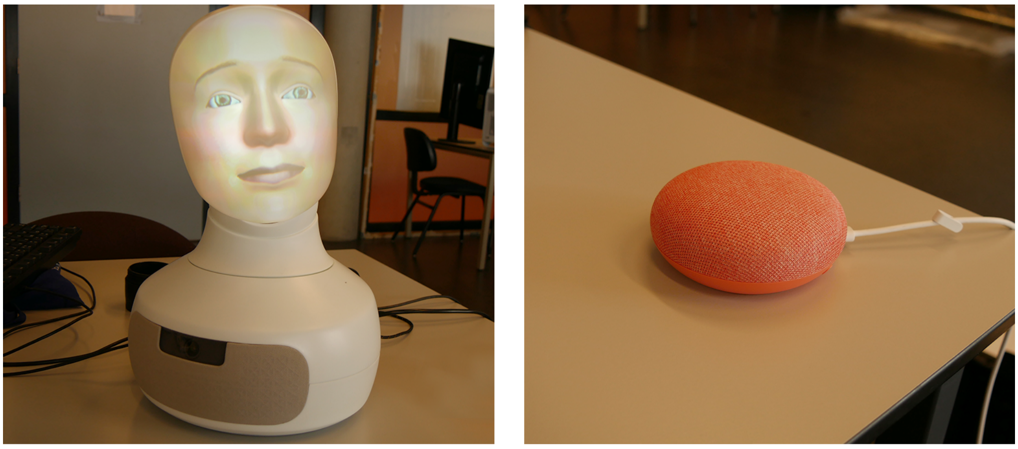
Organisational materials such as parental consent forms were given on paper. The child consent form and post-questionnaire were filled in on the computer, through the survey platform Qualtrics. The thesis website contains all these materials [97].

Lastly, camera equipment was used to record the child's interaction with the agent. The use of recording equipment had several reasons. Firstly, due to privacy concerns, the agents were implemented to not save any information given by the child. Secondly, the researcher was not present during the conversation, to avoid influencing the child. Therefore, watching the video/audio recording was the only option to determine the information disclosure score. The privacy sensitive information that was given in the video was edited out by the researcher after each session, to mediate privacy concerns of parents. Additionally, the video recordings allowed a way to analyse the impact of the agent on the child's emotional state. More on this is explained in the measurements section (see Section 7.8).

7.5 Procedure

Before starting the research, the researcher went to the two primary schools to introduce herself. She went to every class separately and used a PowerPoint presentation to tell the children about herself, her study and what she was planning to do at the school. The exact PowerPoint presentation can be found on the thesis website [97]. The children were told that the goal of the research was to create a "smarter, better robot". Nothing was said about the actual goal of the research. The children were also shown videos of conversations with the agents, to make them familiar with the agents. At the end of the presentation, information brochures and consent forms were handed out for the parents. The researcher went by the schools multiple times in the weeks that followed, which resulted in 86 filled in consent forms.

At the start of each session, the child was briefly informed by the researcher about the session and asked for their consent. After this, the child spoke with one conversational



Furhat (Furhat Robotics)

Google Home Mini (Google LLC)

Figure 7: The embodied conversational agents (ECAs) used for the experiment.

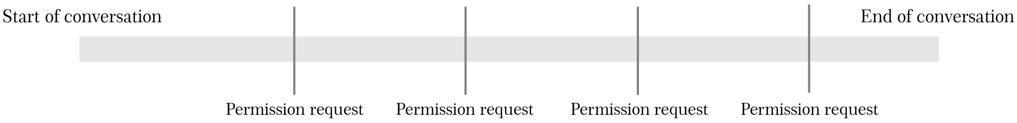


Figure 8: Permission requests throughout the conversation.

agent, either the Google Home Mini or the Furhat robot. There were several moments during the conversation (see Figure 8) where the ECA wanted to collect privacy sensitive information of the child and requested permission to do so. The information that the agent requested, was determined by the pilot study as found in Section 5. The agent either stayed exactly the same during the entire conversation or a stranger presence within the agent occurred during permission requests. Looking at the affordances of the agents' embodiment, this entailed that the Google Home Mini changed its voice for the duration of the request and the Furhat changed its voice as well as its face. More in depth explanation on these changes can be found in Sections 6.4 and 7.6.

After the conversation, the child filled in a post-questionnaire that asked about their perception of the agent and their understanding of the stranger, if present (see Section 7.8). The setup of the experiment can be seen in Figure 9. Both agents were always present in the room. Each child was told specifically with which robot they would talk at the beginning of the session.

7.6 Conversation design

The design principles from Krol et al. for robust experimental design in security and privacy user studies inspired the design of the child-agent conversation [47]. The three main principles they propose are as follows: 1) Give participants a primary task, 2) Ensure participants experience realistic risk, 3) Avoid priming of the participants.

7.6.1 The primary task

Each participant was informed that they would have a five minute conversation with a robot. They were told that mostly the robot will ask them questions, sometimes specifically requesting a "yes", "no" or "maybe" answer, sometimes accepting any answer. It

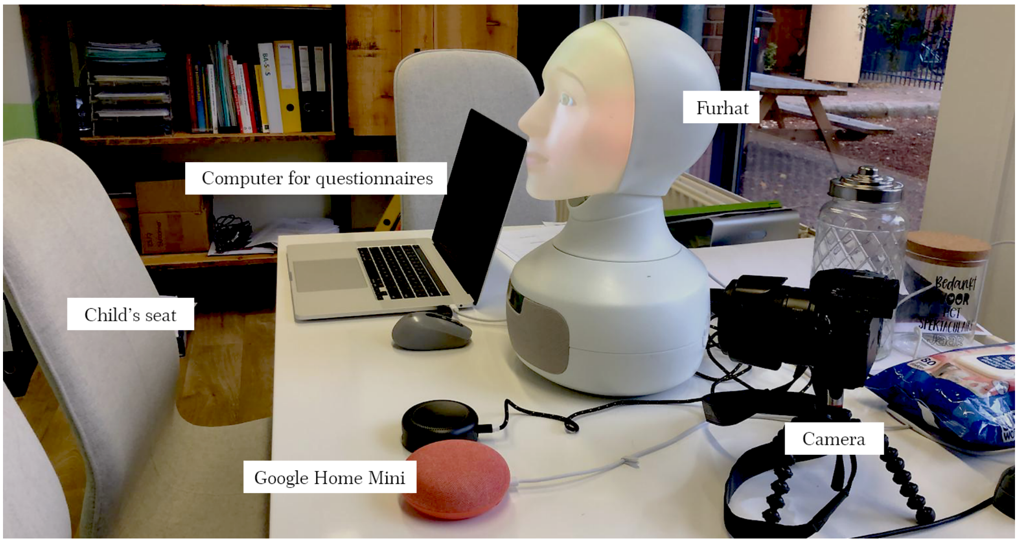


Figure 9: Experiment set up at primary school.

was also said that they are the expert and that there were no wrong or right answers. They were told that they would be recorded during the conversation and that the researcher would not be in the room during that time. She told them that she needed this recording, so she could check how the conversation went. The main objective then, was to see if the robot understood the child and to identify its "weak spots". In other words, finding out what could be improved about the robot. Lastly, each child was told that after the conversation that questions would follow about their perception of the robot and what they thought of the conversation.

7.6.2 Ensure participants experience realistic risk

According to privacy laws, a user has to give explicit consent to a company, when privacy sensitive information is to be collected. This research aims to do the same through privacy permission requests. Usually, these are given through pop-up windows on websites and applications, but an embodied conversational agent can facilitate something similar through audio. For example, the agent in Vitale et al.'s research asked the user for information at different relevant points in the conversation [84]. The only difference with a privacy permission request was that no context was given in that study: the full privacy policy was already provided at the beginning of the conversation with the agent.

The agents in this research do give context at the time of requesting, which resembles a privacy permission request better. There were four moments when the agent asked for privacy sensitive information (see Figure 8). It asked for the child's name, telephone number (from home or their own), computer brand (from home or their own) and access to contacts (two names of people they are closest to). These variables were determined from the first pilot study (see Section 5). Each privacy permission request was neutral in tone and had the same structure:

1. Agent informs the child that it wants to collect a certain variable.
2. It tells them why it needs this information and what will be done with this information.
3. Prompt for a "yes", "no" or "maybe" answer. If "maybe" is said, prompt again for a "yes" or "no" answer.

4. Thank the child for their answer.

In the case of children and robots, information disclosure can happen under pressure of a group of robots for example, such as in Vollmer's research [85]. With these permission requests, it is thought that children will not experience a similar pressure, as there is only one agent and the permission requests will not push towards a correct or wrong answer. No rewards or punishments exist within the experiment either.

7.6.3 Avoid priming of the participants

The background story about improving the robot was told at the start of each session, but also when introducing the research to the children for the first time. Nothing about collecting privacy sensitive information was mentioned. Nothing was said either about a possible stranger appearance. Children in the stranger presence condition got more explanation about the meaning behind the stranger, but only after the child-agent conversation happened.

7.6.4 The child-agent conversation

Regardless of the experiment condition, the child-agent conversation content was always the same to avoid introducing additional factors. The full conversation can be found in Figure 18 in the Appendix. Videos of the conversation can be found on the thesis website [97].

The agent would start with an introduction about itself, calling itself Rob and explaining that he is programmed by a company called Robot B.V. Rob would then ask some questions about the lesson the child was following and offer to make a joke. After this "warm up", that allowed some time for the child to get used to how the ECA worked in terms of listening and responding, the first privacy permission request would occur. In stranger presence conditions, a stranger presence within the agent would now occur. The stranger was supposed to portray Robot B.V., although this was not told to the child until after the conversation. The reasoning behind the stranger being Robot B.V. is explained in the next section.

The permission requests were handled in a very neutral manner. After an answer was collected, the ECA would always change back to "Rob" to further comment on the answer and ask questions related to this topic. Rob gave neutral or positive feedback to the participants, never negative. A privacy permission request would happen three more times in the conversation. After all the requests were handled, Rob would indicate that he had to wrap up the conversation. He would ask the child what they thought of the conversation and wish them well.

7.6.5 Manipulation

In the beginning of the conversation, children are properly introduced to the presence of Rob. Rob tells personal information about himself, cracks a joke and asks the children a few easy questions about their school lessons. This all happens before the stranger presence within the agent takes place, so that children could "warm up" to the presence of Rob and Rob would not be perceived as a stranger. According to Moran et al. [65], knowing a name could already be enough to not view someone as a stranger. Therefore, Rob shares much personal information and many opinions.

When it came to the manipulation, the researcher wanted the children to perceive the stranger presence as a stranger. Therefore, the agent had to be able to change its appearance in order to differ from Rob. Children are often taught about stranger danger in the context of an unknown appearance [4]. The change in appearance was done through auditory and visual changes, based on suggestions from the paper by Major et al. [56]. A change in voice on an Alexa device heightened users' awareness on whether or not they were speaking with a third party or not. This idea was used in this main study as well, as privacy awareness was one of the variables to be researched.

The stranger change within the Google Home Mini device was created through an auditory change. The Furhat robot is capable of changing its visual appearance as well. Therefore, auditory and visual changes were used for the Furhat to represent the stranger appearance. As mentioned in Section 6.4, a stranger does not need to stand out when it comes to features, to be a stranger [4]. The "stranger presence" within the agent was therefore quite "normal", as it should not nudge the child towards certain answers because of its looks. Both ECAs are male agents since the male voices between the two programming platforms were easier to match in tone than the female voices.

Besides the auditory and visual changes, the content of the conversation was also created carefully. For example, the stranger presence never reveals its identity during the conversation. Children are only told about its identity after the conversation. Its given identity was Robot B.V., the company that "created Rob". This identity was chosen because a company usually collects users' information and the people behind a company are usually strangers to its users. Therefore, this identity seemed appropriate and might change children's quick trust when it comes to companies. Besides its identity, the stranger presence also never discloses any other personal information about itself, whereas Rob does share details about himself.

Lastly, talking to a stranger a lot might also lessen the feeling of them being a stranger. Therefore, the stranger presence never says more words or sentences than Rob, so that Rob is the most prominent presence during the conversation with the child.

7.7 Implementation

It was important that the children believed that the agents were autonomous and not tele-operated. Tele-operated agents have been proven to lower information disclosure and heighten feelings of judgement [54]. The researcher was also not present during the conversation, to not impose feelings of judgement onto the children or influence them subconsciously. An autonomous ECA also reflects the social robots that are increasingly found in the wild better.

Several platforms were used to make the two agents autonomous. DialogFlow is a Google Cloud Platform Service for Natural Language Processing. The platform provides a Graphical User Interface (GUI) to program Google Home devices. It builds up conversations using intents. An intent is triggered by an input context. This context typically comes from other intents that have been triggered before, where it is specified as output context. Each intent is trained to recognise a certain input by the participant. Only if the input context is correct and the participant's input matches the training phrases in the intent, the agent gives an answer. This answer is also specified within the intent. There are also fallback intents for when the participant is not understood. Such fallback intents steer the user in giving the correct input, so the correct output by the agent can follow. All the intents together make up the conversation.

In DialogFlow, Speech Synthesis Markup Language (SSML) is used to indicate when a sentence starts and ends and what the prosody details (voice pitch, rate and volume) are. To implement the Furhat, a domain-specific-language (DSL) for Furhat, that is built in the Kotlin programming language, is used. It also works with intents, although their implementation is more flexible than in the DialogFlow GUI. This gives the maker more freedom in coding the conversation and creating clearer file structures. The Furhat code was quicker and easier to maintain and modify than the DialogFlow interface allowed.

However, the Furhat pre-programmed recognition for the Dutch language was non-existent for many variables, such as Dutch numbers, common answers (yes/no/maybe) and names (personal names and computer brands). Many variable examples were provided by the researcher so that the Furhat could be trained. The DialogFlow platform was already trained in recognising a lot of these variables in Dutch, except computer brands, which had to be provided as well.

Besides variable recognition, the agents had to be trained to recognise possible input context. The same training sentences were used for the Furhat as for the Google Home Mini.

General fallback intents of the Furhat also had to be implemented for the Dutch language. Those of the Google Home Mini were tweaked, so they were the same as well.

All in all, the Furhat needed more "guidance" to get to the same standards that were already available for Google Home devices. On the other hand, the code-based approach gives more freedom to the maker, which has many advantages for more complicated conversations.

To avoid conversation complications, such as too many failed attempts at understanding the input, both agents were programmed so that the conversation could not easily fail. This would prevent a decrease in user experience and make sure that the experiment would not take too much time. The conversation is therefore semi-guided to prevent it from going into every direction: the agent took most of the initiative. For most questions, any input by the child is accepted. Only for permission requests, where the agent needed a specific input such as "yes", "no" or "maybe", it would not continue without this answer. The agent is the most strict about this answer, since it is part of the information disclosure score. Without it, this score could not be measured.

Lastly, when the agent got a "yes" to collect specific information, it had three attempts to understand a child's answer, so that it could then use this information in further conversation. If the agent still did not understand the child on their third attempt, it would pretend as if it did understand the child's answer and continue the conversation without the information. This way, the flow of the conversation was not completely disrupted.

7.8 Measurements

The two main dependent variables of this study are information disclosure and privacy awareness. The child's answers during the conversation determined the score on information disclosure. The privacy awareness score was determined by the post-questionnaire. Besides this, there were other measurements, such as the manipulation check, perception of the robot and video responses. They are discussed in this section in "order of appearance", following the experiment set up and post-questionnaire (PQ) sections.

7.8.1 Video responses

The children were recorded in order to determine information disclosure scores. Furthermore, their video recordings were also used to give insight into their emotional state as well, similar to research by Leite et al. [48]. The emotional states of the children are used to aid the motivation behind Hypothesis 1.

Screenshots were taken to show the child's first reaction to experiencing a possible stranger presence within the agent. Each screenshot was taken four seconds into the first privacy permission request. At the moment of the screenshot, the agent has just finished saying: "I'd like to know your name. When I know your name, I can address you personally". The parents of the children gave permission for them to be featured in this report. OpenFace software was used to quantify the emotions visible in the children's faces [7]. The software is capable of recognising facial action units (AU): contraction

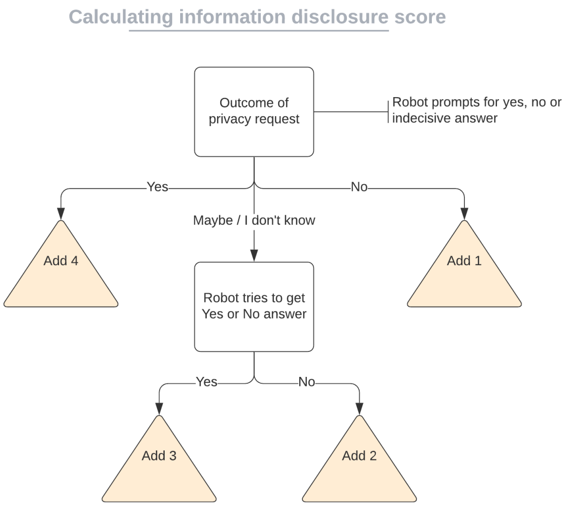

Calculating information disclosure score
Outcome of privacy request
Robot prompts for yes, no or indecisive answer
Yes
No
Maybe / I don't know
Add 4
Add 1
Robot tries to get Yes or No answer
Yes
No
Add 3
Add 2

Figure 10: Information disclosure scoring per question.

or relaxation of muscles at certain places in the face, on a scale of 0 (not present) to 5 (very present). Certain combinations of action units have been found to portray certain emotions [27]. The emotion surprise, as mentioned in the hypothesis, is coded by AU01, AU02, AU05 and AU26. The surprise values of the screenshots are reported in the results of the main study.

7.8.2 Information disclosure

Information disclosure was directly measured from the answers the child gave to the agent. There were four moments in the conversation where the agent requested personal information from the child. Before the information itself was given, children were prompted to give consent first, which required a "yes", "no" or "maybe" as the answer. When "maybe" was given, the agent would ask again for a "yes" or "no". A score was obtained depending on a resounding yes (4), resounding no (1), hesitant yes (3) and hesitant no (2) (see Figure 10). Including "I don't know" as another final answer would have created a five-point Likert scale. However, it is common that privacy permission requests need a definite yes/no answer, otherwise the next state of the entity (such as a website, application or robot) will be unknown. Therefore, the scale only has four options. Besides this, a distinction is made between a resounding and a hesitant answer, because it provides more insight into the child's willingness to give out the information. All variable scores for one child were summed together and divided by four, which led to an overall score on information disclosure.

7.8.3 PQ section 1: General questions

The post-questionnaire started with two questions on participant ID and participant condition, both filled in by the researcher. Age and gender were then asked to the child.

7.8.4 PQ section 2: Manipulation check

The manipulation check served as a way to get insight into children's conceptual model. The goal was to identify whether children perceived the stranger presence as a stranger. To pass the manipulation check, children had to indicate that they noticed a change in the ECA and describe the change (auditory and/or visual change) accurately. Furthermore, they needed to identify the stranger presence as a stranger. This resulted in the following questions:

- Did you notice a change in the robot when it requested permission to collect information? (Yes/No/I don't know)
- What change did you notice? Describe it.
- With whom did you speak when the robot requested permission?

  □ With a robot that was a stranger to me.

  □ With the same robot as from the beginning.

The children who answered some or all of these questions differently, were removed from statistical analyses.

7.8.5 PQ section 3: Information disclosure beliefs

After the manipulation check, the following question was asked to all children, regardless of the condition. It is an exploratory, multiple-choice question and gave additional insight into the conceptual model of children. In other words, how children view the relationship between a product and the product maker: the company. The question was as follows:

- Select the answer you agree with the most:

  □ I gave out personal information to Rob.

  □ I gave out personal information to Robot B.V.

  □ I gave out personal information to Rob and Robot B.V.

  □ I gave out zero personal information.

Comparing the situation to one in the wild, the most probable correct answer is Rob and Robot B.V.. The companies behind conversational agents such as Alexa and Google Home, have access to information given by users. The agents themselves also have access to this information, in order to have personalised conversations with the user.

After this question, children were given additional explanation before moving on to the other questions. Children in the stranger conditions were told that the stranger during the permission requests was supposed to represent Robot B.V.. Children from the no stranger conditions were reminded once more that Robot B.V. is the company behind Rob. This refresher was relevant, as the questions that followed, distinguished between these two entities.

	Privacy Awareness Statements
1	Rob has a male voice
2	I was told what my personal information can be used for
3	The computer brand that we have at home can be used to share information between the robot and our computer
4	Robot B.V. makes robots
5	My name can be used to personally call me by my name
6	My info is saved for 30 days before it gets deleted
7	Rob is funny
8	My contacts can be used to understand better what kind of person I am
9	It is possible to change the information that I have given
10	My phone number can be used to check whether or not the robot has talked to me before

Table 7: Privacy awareness statements from the post-questionnaire, which use a five-point Likert scale from totally incorrect to totally correct (Dutch scale labels by Van Straten et al [82]).

7.8.6 PQ section 4: Privacy awareness

Previous research has often measured privacy awareness by using a memorisation and understanding method. This means that participants would have to answer statements/question about the content of privacy policies [11], [19], [46], [58], [72], [77].

The privacy awareness statements that were asked in this research, are based upon the content of the privacy permission requests of the child-agent conversation. During the requests, the agent gave context for each variable about what this information would be used for. Children's memorisation and understanding of this was tested in statements 2, 3, 5, 8 and 10 (see Table 7). Some other statements that were included, were irrelevant to privacy awareness (1, 4, 7) or were untrue based on the conversation (6, 9). This was done to keep the child alert.

A five-point Likert scale from totally incorrect to totally correct, was used for the statements. The Dutch labels of this Likert scale are taken from research by Van Straten et al [82]. The final score on privacy awareness is the average of scores on statements 2, 3, 5, 8, and 10.

7.8.7 PQ section 5: Perception of the agent

Lastly, the questionnaire included items to determine the perceived likeability, intelligence and trustworthiness of the agent. These variables are often asked in CRI research [23], [48], [91]. While the research question does not directly require these measurements, they provided additional insight into the bond that the child created with the agent, that was used to motivate hypotheses. The stronger the bond that a child creates with the agent, the more likely they might be to give out personal information to the agent. This makes these variables relevant to research.

The likeability and intelligence items as asked were taken from the translated version [83] of the godspeed questionnaire developed by Bartneck et al [9]. Children had to indicate their position towards the agent "as a whole" on a semantic differential scale between two bipolar words. All opposites can be seen in Table 8. Children in the stranger conditions were told specifically to answer these questions with both Rob and Robot B.V. (the stranger) in mind. Therefore, with a large number of participants, the differences between stranger and no stranger conditions should be explainable by the presence or absence of a stranger.

Furthermore, the trust the child had in the agent was determined by statements taken from CRI research by Van Straten et al [82]. The trust statements were given for both Rob and Robot B.V., to see whether children within each condition trusted Rob and the company behind Rob differently (see Table 9). Since children easily trust companies up until the age of 12, it is insightful to see whether similar results for Rob and Robot B.V. will be found in this research.

	Likeability	Intelligence
1	Dislike - Like	Incompetent - Competent
2	Unfriendly - Friendly	Ignorant - Knowledgeable
3	Unkind - Kind	Irresponsible - Responsible
4	Unpleasant - Pleasant	Unintelligent - Intelligent
5	Awful - Nice	Foolish - Sensible

Table 8: Agent perception items from the post-questionnaire, given with a five-point semantic differential scale [9].

	Trust Questions
1	I feel that I can trust X
2	I feel that X can keep one of my secrets
3	I feel that X is honest
4	I feel that X is trustworthy

Table 9: Statements on trust from the post-questionnaire [82], given with a five-point scale from totally incorrect to totally correct. X was replaced with Rob and Robot B.V., which resulted in 8 statements.

7.9 Results

In this section, the results of the main study experiment are presented. The statistical tests are performed on the 60 remaining participants (see Table 6 for participant details). Firstly, the effects of the independent variables (a stranger (non)presence and level of embodiment) on information disclosure are addressed. Secondly, the effects of the independent variables are looked at for the privacy awareness scores. The correlation between the two dependent variables (information disclosure and privacy awareness) are reported after this. Lastly, other measurements that have been collected through videos and the post-questionnaire are reported.

7.9.1 Information disclosure results

As mentioned in Section 7.8 on measurements, the scores that were collected on information disclosure ranged from 1 to 4. This means that when a mean score is 1, a participant always gave a resounding no and thus disclosed no data. When a mean score is 4, a participant always gave a resounding yes and thus disclosed all their data. Information disclosure is looked at both in regards to stranger (non)presence and the interaction between stranger (non)presence and the level of embodiment. Because information disclosure is looked at for multiple factors, a two-way ANOVA would be a suitable test to perform, if all assumptions are met.

Checking assumptions Before a two-way ANOVA can be performed, six assumptions need to be passed. The first three assumptions are passed because of the experiment design: the information disclosure variable is measured at a continuous level, the two independent variables both contain two or more groups and there is independence of observations (all conditions stand separate). The last three assumptions are checked using SPSS software. Two out of three were passed: there were no outliers found in the data and all combinations of groups (conditions) passed the Levene's test, which meant that there was homogeneity of variances. Lastly, the data needed to be normally distributed for each combination of groups. This can be analysed using the Shapiro-Wilk test.

Testing for normal distribution Firstly, as H1A focuses on the effects of a stranger presence, we check normality for the NS and S conditions. The average score of the NS condition (N=39) is $M = 2.4487$ ($SD = 0.70520$), the average score of the S conditions (N=21) is $M = 2.2143$ ($SD = 0.64365$). It was found that the distribution of both NS and S conditions were skewed towards the left side. This meant that overall, more participants scored lower on information disclosure than higher. Furthermore, to pass a normality test, the p-values on the Shapiro-Wilk test need to be higher than $p = 0.05$. NS and S values were $p = 0.015$ and $p = 0.041$ respectively, which meant that they did not pass the test for normality. Transforming a variable using log can sometimes help to achieve normality: therefore, a log transformation was performed on the information disclosure variable, after which the normality test was performed again. Both conditions still did not pass the test (NS: $p = 0.046$, S: $p = 0.044$).

Whether the conditions F and GH were also normally distributed, was checked next. The average score of the GH condition (N=26) was $M = 2.3750$ ($SD = 0.6334$), the average score of the F condition (N=34) was $M = 2.3603$ ($SD = 0.7365$). The GH and F p-values on the Shapiro-Wilk test were $p = 0.002$ and $p = 0.046$ respectively, which meant that they did not pass the tests for normality. The GH scores were skewed towards the left side. The log transformation did not help to pass the tests for normality either (GH: $p = 0.013$, F: $p = 0.024$).

	Information disclosure
F NS	2.5000 (0.76376)
F S	2.1833 (0.68444)
GH NS	2.4000 (0.66094)
GH S	2.2917 (0.57915)

Table 10: Information disclosure means (standard deviations) per condition (1: lowest mean value, 4: highest mean value).

As H1B focuses on an interaction effect between stranger presence and level of embodiment, normality is also checked for all conditions (GH NS (N=20), GH S (N=6), F NS (N=19) and F S (N=15)). The separate mean scores of all conditions can be found in Table 10. The Shapiro-Wilk test was performed to check for normality. One out of four conditions did not pass this test: the GH NS condition was not normally distributed ($p = 0.005$) as it was skewed towards the left side. A log transformation was performed on the information disclosure scores. It brought the p-value of the GH NS condition up to $p = 0.028$, yet it stayed significant, which meant a normal distribution of the data could not be assumed.

Used test Since not all assumptions were met in order to use a parametric test, a non-parametric test was needed. A non-parametric equivalent to the two-way ANOVA was performed, using the raov function of the Rfit package, developed by Kloke and McKean [32]. The original information disclosure scores were used for this test.

Effects of stranger (non)presence A p-value of $p = 0.380$ ($F(1, 56) = 0.785$) was found for the effects of a stranger presence, meaning that the differences between the NS and S conditions were insignificant, since they were bigger than $p = 0.05$. This means that hypothesis H1A, that a stranger non presence leads to more information disclosure, cannot be accepted.

Interaction effect between stranger (non)presence and level of embodiment An interaction effect was hypothesised to occur between the two independent variables. In Section 4.1, a Figure on the hypothesised interaction can be found. It was thought that children would disclose more information to a “no stranger presence” within a high embodied agent, but less to a “stranger presence” within a high embodied agent, compared to a low embodied agent. The mean values as were found in the main study (see Figure 11) show a similar pattern to the hypothesised figure. However, the p-value of the interaction between embodiment of the agent and stranger (non)presence was $p = 0.751$. This means that hypothesis H1B cannot be accepted, as the differences between the conditions are not significant.

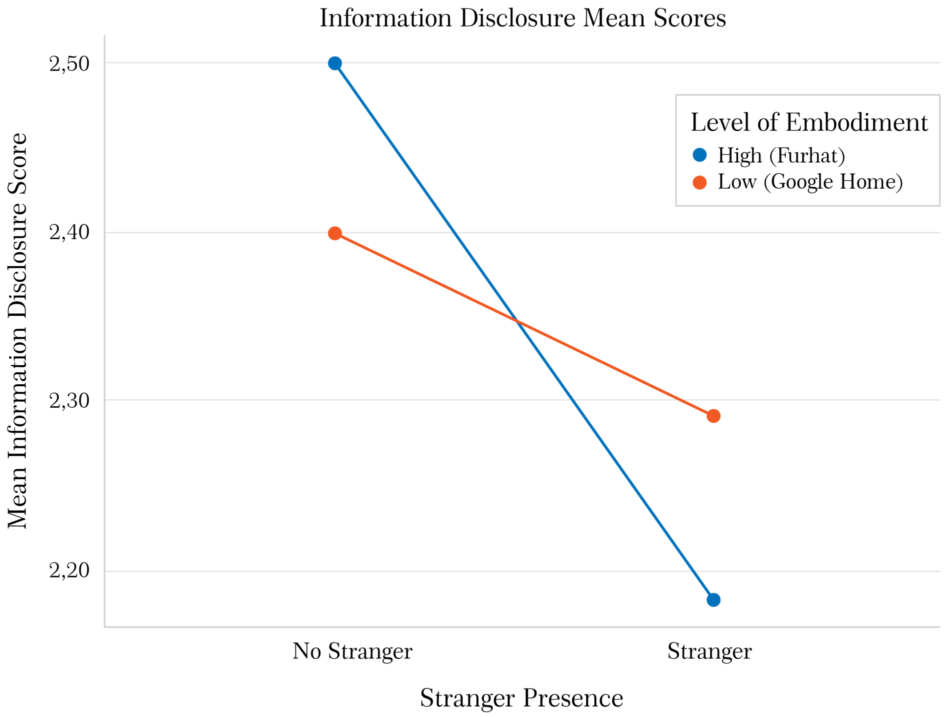


Figure 11: Information disclosure means per condition. Please note the zoomed in scale of the Y-axis.

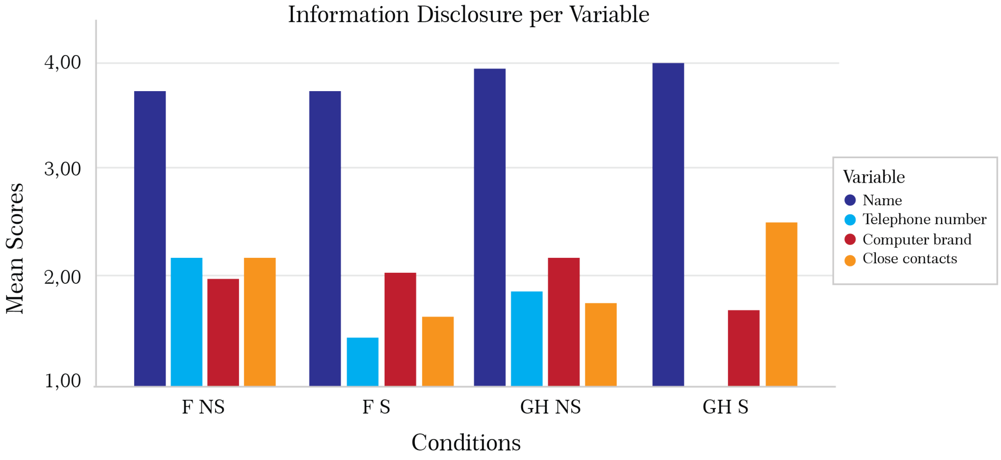


Figure 12: Information disclosure means per variable per condition.

	Pilot study	F NS	F S	GH NS	GH S
Name	2.27 (0.961)	3.74 (0.733)	3.73 (0.799)	3.95 (0.224)	4.00 (0.000)
Phone number	1.55 (0.710)	2.16 (1.302)	1.40 (0.910)	1.80 (1.105)	1.00 (0.000)
Computer info	1.30 (0.560)	1.95 (1.311)	2.00 (1.254)	2.15 (1.268)	1.67 (1.211)
Contact access	1.12 (0.288)	2.16 (1.167)	1.60 (1.242)	1.70 (1.031)	2.50 (1.643)
	1.56 (0.476)	2.50 (0.764)	2.18 (0.684)	2.40 (0.661)	2.29 (0.579)

Table 11: Information disclosure means per variable (1: lowest mean value, 4: highest mean value).

	Privacy awareness
F NS	3.9684 (0.57063)
F S	3.7200 (0.47689)
GH NS	3.9300 (0.54396)
GH S	4.0667 (0.56095)

Table 12: Privacy awareness means (standard deviations) per condition (1: minimum average, lowest privacy awareness, 5: maximum average, highest privacy awareness).

7.9.2 Privacy awareness results

The average privacy awareness scores can range from 1 (minimum average, lowest privacy awareness) to 5 (maximum average, highest privacy awareness). Both level of embodiment and stranger (non)presence were analysed in regards to the privacy awareness scores. The mean scores for all conditions are reported in Table 12, and visually presented in Figure 13. This figure is given since the hypotheses in Section 4.2 made a prediction on its general shape. It can already be noticed that the two do not match, mainly caused by the found mean score for the F S condition.

Checking assumptions Again, as multiple factors are at play here, it is appropriate to use a two-way ANOVA. The main effects of the agent's embodiment and a stranger (non)presence were of interest, since H2A and H2B made predictions on this. First, it was checked whether all assumptions were met in order to run the two-way ANOVA. The dependent variable is continuous, the two independent variables consist of two or more categorical, independent groups and there is an independence of observations. As for statistical tests, no outliers were found in the data. Furthermore, all combinations of groups passed the Shapiro-Wilk tests on normality. Lastly, all combinations of groups passed the Levene's test, that tested if there was homogeneity of variances. This was the case. Because all these conditions were met, a two-way ANOVA could be used.

Effects of level of embodiment The averages of the embodiment conditions are $M = 3.8588$ ($SD = 0.53830$) for the F conditions and $M = 3.9615$ ($SD = 0.53969$) for GH conditions. It was hypothesised that children would show higher privacy awareness when an ECA has a high level of embodiment (the F condition). This was not the case, since the main effect for level of embodiment yields the following results: $F(1,56) = 0.975, p = 0.328$. The p-value is insignificant, hence, hypothesis H2A cannot be accepted.

Effects of stranger (non)presence The average privacy awareness score over the NS conditions is $M = 3.9487$ ($SD = 0.55006$), the average over the S conditions is $M = 3.8190$ ($SD = 0.51344$). It was hypothesised that children would show higher privacy awareness when an ECA has a "stranger presence" (the S condition). The main effect for stranger (non)presence resulted in $F(1,56) = 0.128, p = 0.722$, indicating no

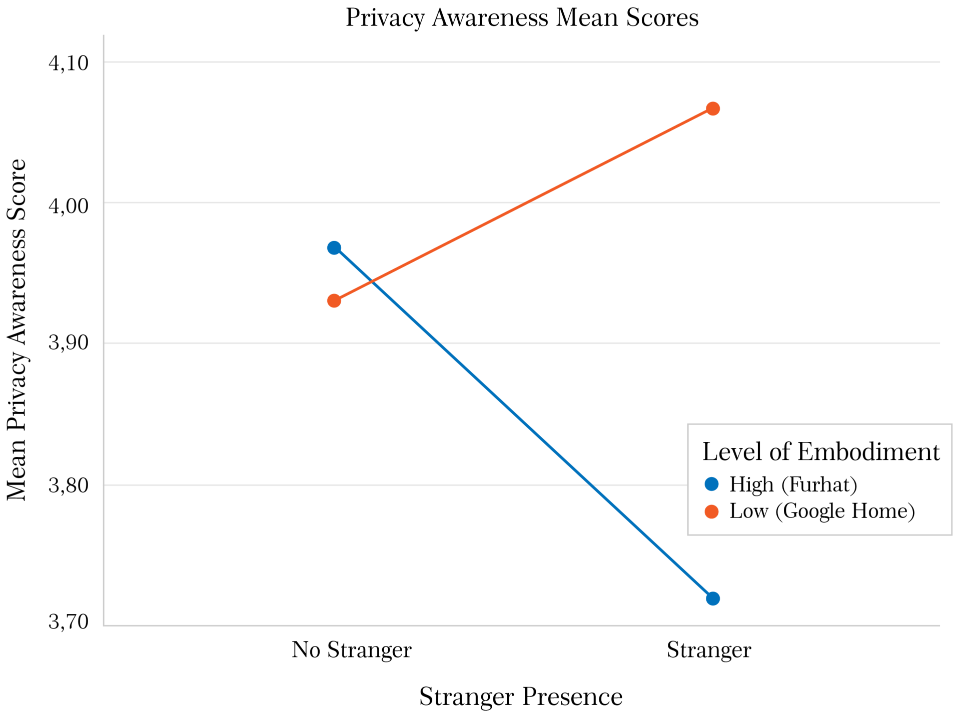


Figure 13: Privacy awareness means per condition. Please note the zoomed in scale of the Y-axis.

significant differences between the stranger and no stranger condition in regards to the privacy awareness scores. Hypothesis H2B can therefore not be accepted.

7.9.3 Correlation results

It was hypothesised that a negative correlation would occur between the privacy awareness scores and information disclosure scores. It was thought that higher information disclosure scores would relate to low privacy awareness scores and the opposite way around. In order to analyse this, a Pearson correlation coefficient was computed to assess this relationship. There was no correlation found between the dependent variables, as the Pearson correlation indicated a value of $r = -0.036$. A value of (-)1 means that there is a strong correlation, a value around 0 indicates no correlation. A scatterplot was created (see Figure 14) to show the relationship trend line. As the found correlation is so low, the points do not show any linear trend and appear to be "scattered".

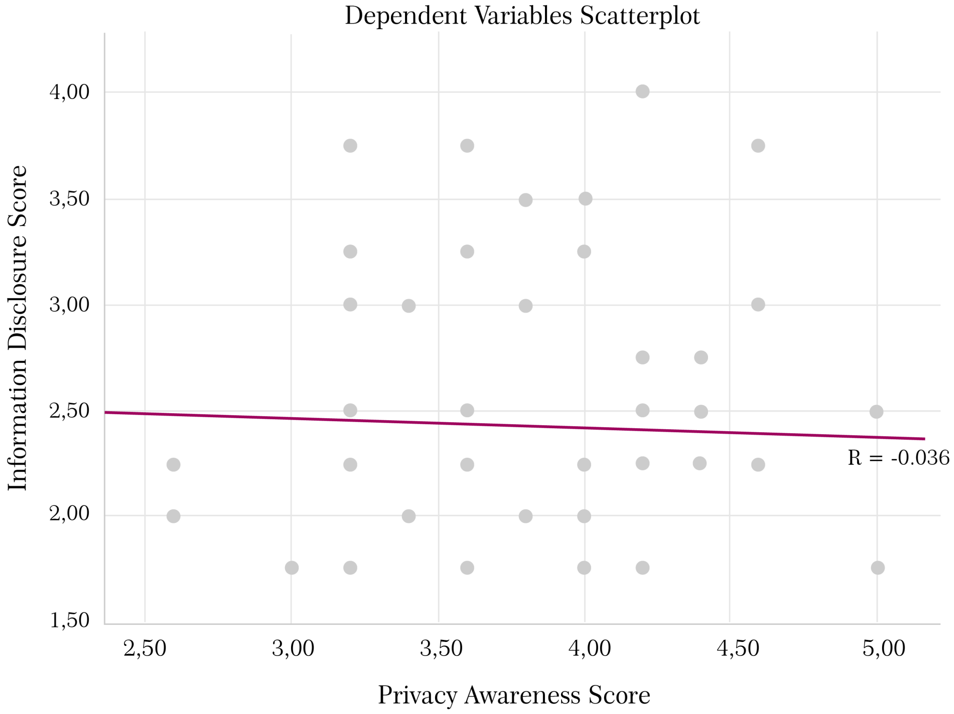


Figure 14: Scatterplot on information disclosure and privacy awareness scores.

7.9.4 Other measurements

No statistical tests were performed on the measurements in this subsection, but they are used as motivation for Hypothesis 1. Hypothesis 1 mentions the bond that the child creates with the agent.

Video responses Firstly, from the video responses, it was possible to extract Action Unit mean values. Certain combinations of Action Units make up emotions. The emotion "surprise" is mentioned in Hypothesis 1 in an argument that a sudden change into a stranger presence would lower information disclosure scores, partly because of the negative affect that it could bring. A grid of images was created to show the reactions of children to the first privacy request made by the agent (see Figure 15). The AU mean strength values (on a scale of 0 to 5, with 5 being the highest intensity score) can be found in Table 13. The facial muscles, that are represented by each action unit, can be found in Table 14. All action units lie between 0 and 1.175. AU01, AU02, AU05 and AU26 make up the emotion surprise. There are no clear differences between the conditions that indicate that S conditions score higher on this emotion, as S conditions do not consistently score higher on all the surprise AUs compared to the NS conditions (see Table 13).

	AU01	AU02	AU05	AU26
F NS	0.228	0.210	0.725	0.700
F S	0.733	0.080	0.140	0.545
GH NS	0.842	0.000	0.000	0.953
GH S	0.293	0.000	0.000	1.175

Table 13: Action unit mean values as calculated from the screenshots of Figure 15, using OpenFace software [7].

Action Unit	Description
AU01	Inner Brow Raiser
AU02	Outer Brow Raiser
AU05	Upper Lid Raiser
AU26	Jaw drop

Table 14: Action units and their corresponding facial muscle movement.

Perception of the agent Besides the video recordings, measurements were obtained through a post-questionnaire. The children filled in statements on the likeability of the agent, the intelligence of the agent, trust in Robot B.V. and trust in Rob. The means of these variables are reported in Table 15, to provide insight into the thoughts of the children. Due to the nature of the scales, the most positive average is 5, the most negative average is 1. In Figure 16, a visual representation of the mean values is plotted. Overall, children really liked each agent ($M = 4.2533$), found it intelligent ($M = 4.2400$) and trusted it (Rob: $M = 3.9333$, Robot B.V.: $M = 3.7875$). All scores are quite high, around four. Therefore, it may be assumed that children were able to bond with the agent in each condition.

Children's conceptions Lastly, the results of the exploratory question, to whom children believed they gave their data, are summarised in Table 16. Figure 17 shows a stacked bar plot for these results. This plot shows more variation between conditions than the perception of the agent plot did. As can be noticed from the figure, the GH NS condition differs the most from other conditions. 55% of children in the GH NS condition believed they were only speaking to Rob, while children in the F NS condition more often thought that Robot B.V. also played a role (42.1%), even though there was no stranger presence. Furthermore, for both S conditions counted that more children thought that they gave out no information at all, compared to the NS conditions (F NS: 21.1%, F S: 33.3% ; GH NS: 25.0%, GH S: 33.3%). They also more often thought that they only gave info to Robot B.V. (F NS: 10.5%, F S 20.0% ; GH NS: 0%, GH S: 16.7%). This was before it was revealed to them that Robot B.V. represented the stranger presence. The discussion (Section 7.10) will go into further depth on these results.

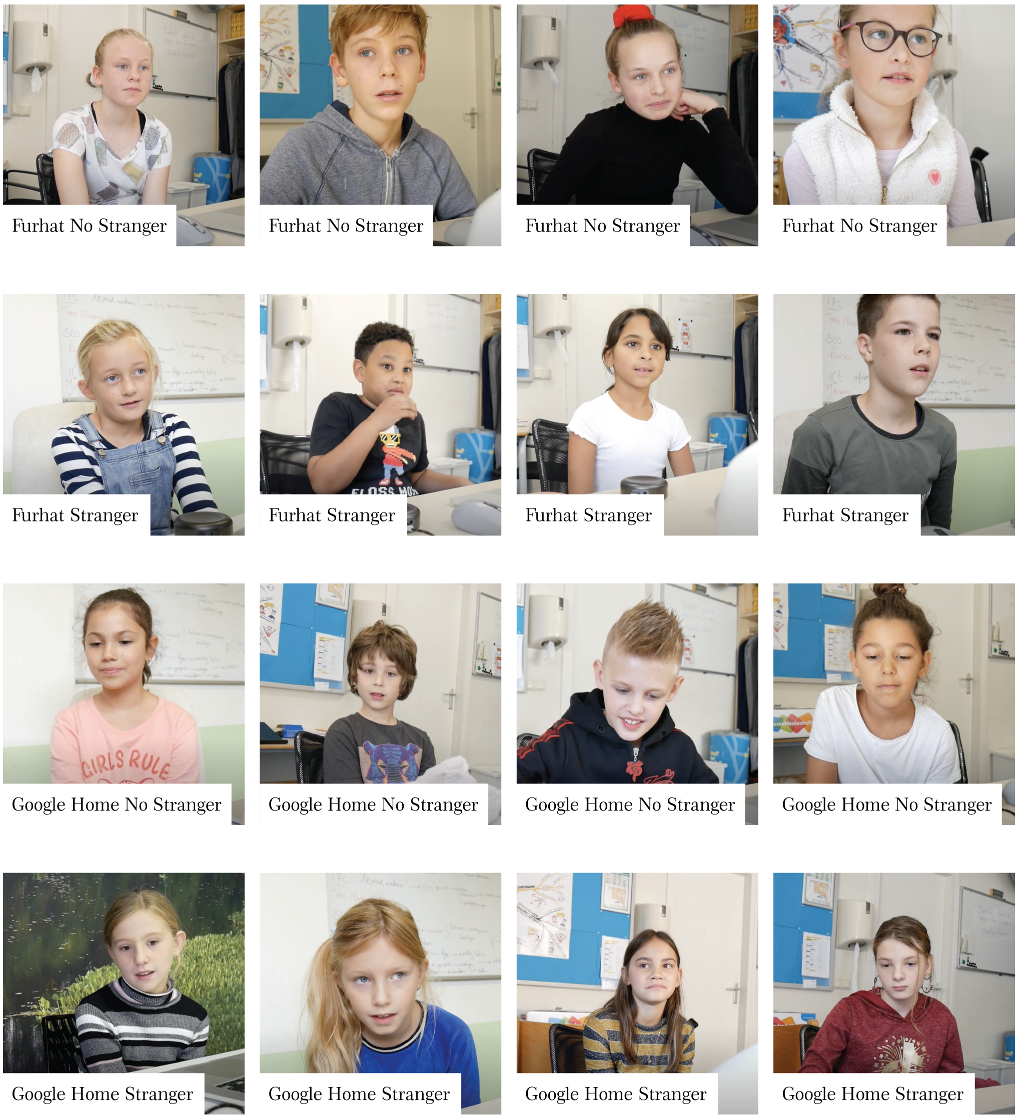


Figure 15: Sample screenshots of children's facial expressions during the first privacy permission request, each row containing one of four conditions. Top to bottom: F NS, F S, GH NS, GH S.

	Likeability	Intelligence	Rob trust	Robot B.V. trust
F NS	4.2316 (0.55882)	4.2316 (0.46314)	3.9342 (0.84097)	3.9605 (0.63060)
F S	4.1200 (0.64940)	4.3200 (0.47689)	3.9333 (0.49522)	3.6833 (0.72866)
GH NS	4.2500 (0.65172)	4.1500 (0.61857)	3.8125 (0.96612)	3.5875 (0.85945)
GH S	4.6667 (0.53166)	4.3667 (0.55737)	4.3333 (0.51640)	4.1667 (0.49160)
Overall	4.2533 (0.61492)	4.2400 (0.52439)	3.9333 (0.78636)	3.7875 (0.73908)

Table 15: Mean values (standard deviations) of children's perception of the agent per condition.

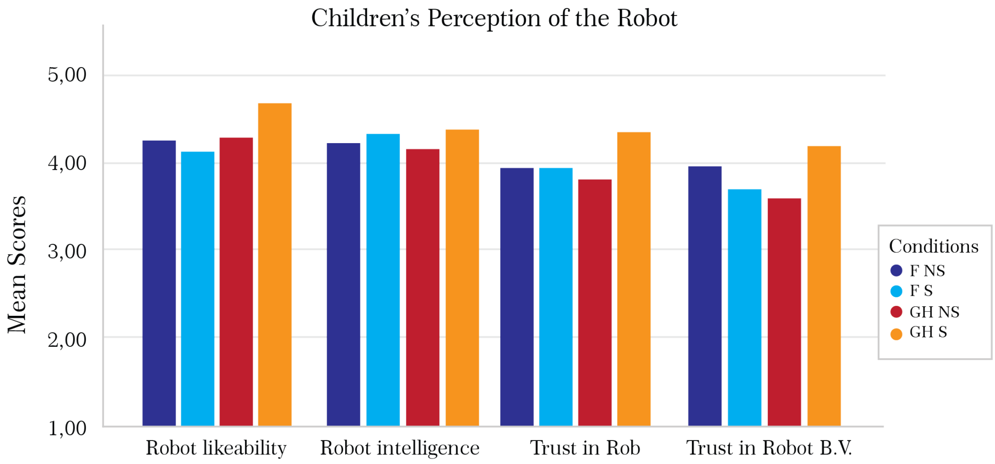


Figure 16: Mean values of children's perception of the agent per condition.

	None given	Only Rob	Only Robot B.V.	Both
F NS	21.1%	26.3%	10.5%	42.1%
F S	33.3%	20.0%	20.0%	26.7%
GH NS	25.0%	55.0%	0.0%	20.0%
GH S	33.3%	33.3%	16.7%	16.7%
Overall	26.7%	35.0%	10.0%	28.3%

Table 16: Initial thoughts of children on who they gave their personal information to.

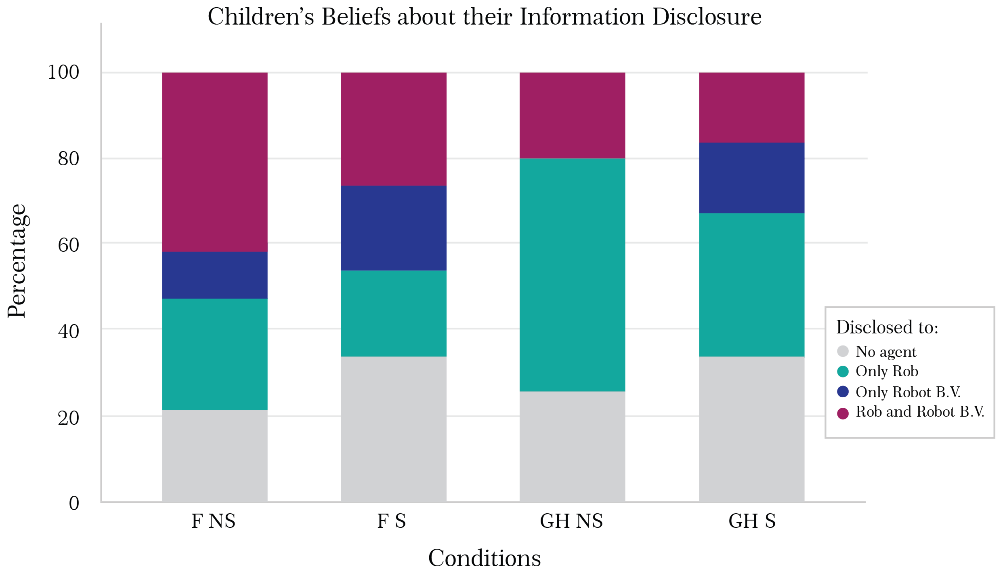


Figure 17: Initial thoughts of children on who they gave their personal information to.

7.10 Discussion

The section first delves into the results that are connected to the hypotheses of this research. Post-questionnaire measurements as well as observations and children's own comments and beliefs are used to provide more insight into the results.

7.10.1 Information disclosure

Hypothesis 1A It was hypothesised that children would disclose more information to a "no stranger presence" compared to a "stranger presence", within an embodied conversational agent. The analysis showed that there are no significant results to prove this. It was thought that a stronger bond would be created in the NS conditions, because the child would have spent more time with the same agent, Rob. The children's likeability scores can be looked at for an indication of this bond. As can be seen in Figure 15 and Table 16, children in each condition indicated that they liked the overall agent very much. There are no clear differences between agents. It is assumed that children were able to bond with the agent in every condition. The novelty of talking to a conversational agent might have played a role in this, as many children had never spoken to a Google Home or Furhat before and were very excited about this.

Something similar happened in Leite's research, where children indicated that they liked the robot very much, even though parents saw negative emotions in their faces when the robot revealed a secret that it was not supposed to know [48]. Therefore, children's facial expressions were also looked at: screenshots were made of the first crucial privacy permission request, with children from each condition. These pictures can be found in Figure 15. It was hypothesised that children in the stranger conditions would be surprised by the unexpected event and be less willing to give out information. The researcher sees no extreme differences in facial expressions. From a subjective point of view, the middle two of the four children in the F S condition seem quite surprised, judging by their upright posture and wide eyes. However, the average AU-scores on surprise (coded by AU01, AU02, AU05 and AU26) for the F S condition do not seem to back up this finding, as they do not surpass the values of other conditions on all action units (see Table 13).

The last argument of the hypothesis was that children would disclose less information because they'd perceive the stranger presence as a "threat". Some children seemed to feel this threat and indicated after the conversation that they did not want to give information to the agent: "It (Furhat) really was a different person! I didn't trust him. I liked the other robot a lot better." (F S, M, 10). Others did not feel threatened: "I thought it was fun (the conversation), but why did you (Google Home Mini) want to know so much personal information?" (GH S, M, 10). The opinions that children shared thus varied on this topic.

Hypothesis 1B For the NS conditions, it was hypothesised that a higher embodiment would facilitate more information disclosure. There were no significant results to prove this. Observations show that personal preference might have played a role in this. Many children expressed an opinion on the Furhat on entrance, that varied from negative ("he looks too human", "he is scary) to positive ("he looks really cool"), more so than on the Google Home Mini. A lot of boys stated a preference to talk to the Furhat robot over the Google Home Mini. There were thus varied personal preferences within and between each agent. That no clear differences in information disclosure were found for the different levels of embodiment, could be because of these personal preferences.

In the S conditions, an opposite effect was expected. In this case, a high embodied agent was thought to match better with children's expectations of a stranger, thus making a bigger impact: this would then lead less information disclosure. Although the results are insignificant, there are some comments made by children that seem to indicate that the Furhat robot had a bigger impact: “It (Furhat) really was a different person! I didn’t trust him. I liked the other robot a lot better.” (M, 10), “I got scared because of his (Furhat’s) face, because it was suddenly so different.” (F, 8) and “It (Furhat) was quite scary, especially the first time I saw the change. Each time after that I got more used to it.” (M, 10). One child even left the room during the first privacy request to get the researcher, because she found the “stranger” scary. No comments with this much affect were said about the Google Home Mini. One participant (F, 8) said that “When it (Google Home Mini) switches to a different voice, I don’t know very well with whom I’m talking, who that is,” but nobody indicated in words that they were scared in the GH S condition. Even though the stranger presence in the F S condition seemed to have a bigger impact than the GH S condition, it did not lead to significantly different information disclosure scores.

Information disclosure per variable Only the average information disclosure score was used for the statistical analysis. However, the separate scores for each variable were reported as well, to obtain further understanding on which variables children in the main study found privacy sensitive. The scores are reported in Table 11 and Figure 12. Looking at Figure 12, it is noticeable that children gave out their phone number and contacts less in the F S condition, compared to the F NS condition. In the GH conditions, children gave out their phone number and computer brand less in the GH S condition, compared to the GH NS condition. Children did give their contact access more often in the GH S condition compared to the GH NS condition. As each variable that was asked during the conversation was more privacy sensitive (based on the first pilot study), it might be expected that information disclosure scores would lower with each question. The Figure shows that this was not the case.

It should be mentioned that the phone number question introduced some confusion amongst children. While the agent explicitly said that the number could be their own or their parents’, children often said they did not have a personal phone or that they did not know their personal number by heart, ignoring the parental option in the question. Multiple children (in all conditions) shared that, if they had known a number, they would have given it. The type of computer brand at home was another variable that children sometimes did not know. Again, some indicated that they would have provided this if they did know it. This occurred less often than the telephone number. The information disclosure scores are, most probably, lower because of this.

Children’s information disclosure beliefs To get additional insight into children’s conceptual models, they were asked directly after the conversation on who they thought they gave out their personal information to, even before the children in the S conditions were told that the stranger was supposed to represent Robot B.V.. So at the time of asking, the only information all children knew about Robot B.V., was that Robot B.V. created Rob. Table 16 and Figure 17 show their answers.

As can be seen from Figure 17, more children thought that they gave out information to only Robot B.V. in the S conditions than in the NS conditions (so F S compared to F NS, and GH S compared to GH NS). Therefore, it might be assumed that the stranger presence led more children to believe that they gave out information to the company Robot B.V.. Children also more often thought that they gave out no information in the S conditions, compared to the NS conditions.

The scores in the F NS and GH NS conditions are also interesting, as many more children in the F NS conditions believe that they gave out information during the requests to both Rob and Robot B.V or only Robot B.V.. Children in the GH NS condition thought a lot more often that they gave information to Rob only, compared to the F

NS condition. The embodiment of the robot may perhaps be an explanation for these differences. Perhaps it is easier to imagine that there is a company behind a more “high-tech” product, such as the Furhat robot.

7.10.2 Privacy awareness

Hypothesis 2A It was hypothesised that children would show higher privacy awareness when an agent has a high embodiment, regardless of a stranger (non)presence. However, no significant differences were found. One argument used in the hypothesis was that children are taught to listen to other humans and the high embodied agent resembled a human more. Children did regard the high embodied agent as more human-looking. Some even thought that it looked “too human” and were a bit afraid of it. The Google Home Mini got described as an “orange bakpao”, “muffin” and “flat thing”, but never as “human like”.

Another reason that was used to explain the hypothesis, was the attention of a child: a high embodied agent might facilitate more attention, leading to a higher privacy awareness. Its visually interesting form and tracking (following the child’s movements) were named as reasons for this attention. Through the video recordings, it became clear that children were certainly intrigued by the Furhat robot. They would even move on purpose during the conversation to see whether the robot would follow them. Possibly, their excitement and curiosity about the agent might even have distracted them.

It should also be mentioned that a few children in the GH conditions were intrigued with the Furhat robot and out of curiosity, sometimes looked at this agent instead of the Google Home Mini that they were conversing with. Still, the privacy awareness scores were high for F and GH conditions, so most information told by the robot was remembered correctly by the children.

Hypothesis 2B It was thought that children would be more privacy aware when an ECA has a stranger presence, regardless of its embodiment. It was thought that they would pay more attention to the presence in order to minimise stranger danger, resulting in higher privacy awareness. There are no significant results to support this. In general, the mean values on privacy awareness are quite high for NS and S conditions, indicating that all children were generally privacy aware. Children did notice the stranger presence, but it cannot be said whether or not they paid more attention to the privacy permission requests in the S conditions than the NS conditions because of it.

The non-difference might be a result of the experiment design. The privacy awareness statements were asked shortly after the conversation ended: immediate recall usually shows better scores than when there is a longer period of time before recall. If the stranger presence facilitated more attention, perhaps that the information would be remembered better in the S conditions after a longer period of time has passed. A longer period of time between the conversation and the post-questionnaire might have given different results on children’s levels of attention, and consequently privacy awareness, in the NS and S conditions.

7.10.3 Correlation between privacy awareness and information disclosure

Hypothesis 3 A negative correlation between privacy awareness and information disclosure was expected, which was build upon the expected differences between the NS and S conditions. The stranger presence conditions were thought to lead to low information disclosure and high privacy awareness, in contrast to the “no stranger presence” conditions. However, no such differences were found and there was also no correlation between the two dependent variables.

As can be seen from the scatterplot (see Figure 14), there are a lot of different combinations of privacy awareness and information disclosure scores, which results in no correlation. It can therefore not be said that a higher privacy awareness score leads to a lower information disclosure score and the opposite way around. It is likely that more reasons than those mentioned in the hypotheses are at play here, such as the background knowledge of a child on online privacy.

Whereas some children freely gave out their information and even wanted to tell the agent additional information about themselves (real examples include: being bullied at school or insecurity about their medical diagnosis), others questioned why the robot even wanted to know this much information in the first place. Children commented: "The first time I talk to somebody, I never trust them immediately." (GH S, F, 8) and "I don't tell secrets to people I don't know so why would I tell this to a robot." (F NS, F, 9).

Furthermore, children had varied ideas about the agents and their capabilities, which might have played a role in their answers. Some children expressed critical views, while others were more naive in their opinions. A few examples of naive comments are: "Robots always put information in a secret file so they can always be trusted" (GH S, M, 9), "Even though he changed, I can become friends with him" (GH S, M, 9) and "Every robot is friendly and can be trusted" (F NS, F, 8). Some critical comments were: "I think the robot is going to send everything to the company afterwards" (F NS, M, 11) and "I think that a robot always sends information to the company, because the company has made the robot" (GH NS, M, 10). It occurred more often that children did not voice critical opinions about the company, Robot B.V., which can be seen back in high trust scores over all conditions (see Table 15). This matches Livingstone's findings, that state that 5-11 year olds are very trusting of companies [53].

Multiple reasons have previously been mentioned for the results on privacy awareness and information disclosure, respectively. The differences between S and NS groups did not result in a clear correlation: other personal factors, such as background knowledge of a child, probably played a role in forming the points as found in the scatterplot.

7.11 Limitations and future research

It is important to acknowledge that the main study does not come without limitations. These are described in the sections below. Suggestions for future research are mentioned as well.

7.11.1 Group sizes and power analysis

In between-subject research, a group size of N=20 is seen as a golden standard [28]. The sample size of the GH S condition (N=6) was small compared to the other conditions (see Table 6). Even though the voice differences between Rob and the stranger presence were made more obvious after the second pilot study, many children in the GH S condition still did not pass the manipulation check. 13 out of 20 children in the GH S condition identified a change in voice correctly, but only 6 of these children identified the presence as a stranger as well. Therefore, the researcher believes that a bigger voice change might benefit the recognition of a stranger. Yet, it could also be that only an auditory change does not easily cause the feeling of a stranger. Educational materials more often focus on visual or visual and auditory examples with stranger danger than only auditory examples [4].

The F S condition also contains less than 20 participants, namely 15 participants. 15 out of 21 participants passed the manipulation check, which can be seen as satisfactory. The group numbers of the GH NS and F NS conditions are up to par, with N=20 and N=19, respectively.

	Post hoc power
H1A	≈ 0.179
H1B	≈ 0.080
H2A	0.163
H2B	0.064

Table 17: Post hoc power for H1 and H2 results, with values ranging between 0 and 1. A power of 0.8 should be strived for. It should be noted that the post hoc power for H1 was created using a parametric test instead of the non-parametric test, which did not provide enough information to do so. The actual power values for H1 are probably lower than these values, as non parametric tests are less powerful than parametric tests.

For future research, it is advised to use a different enrolment ratio: in other words, classifying more children into conditions that need to pass a manipulation check. This would be a good safety measure to reduce group imbalances. More participants in general is also advised.

Lastly, a post-hoc power (PHP) analysis reveals that the power of the performed tests was really low (see Table 17). Group imbalances and sizes play a role in this, but the small mean differences most likely play the largest role in this.

7.11.2 Manipulation

To get through the manipulation check, children had to recognise a change within the agent, name the kind of change and believe that it was a stranger presence. Not every participant got through the manipulation check: these numbers are mentioned in the previous section. Most noticeably, the manipulation check was passed significantly less by those in the GH S condition. While a change was noticed by 13 participants, the presence only felt like a stranger to 6 out of 20 participants. This may perhaps be explained by "stranger danger" education: not many children are taught examples where a stranger is only heard [4]. There is no real emphasis on speech only.

Furthermore, it is the question whether or not a stranger presence was a good approach to lower information disclosure and heighten privacy awareness. As mentioned previously, no significant results were found for the two dependent variables. Besides this, the overall agent was rated very high on likeability in the S conditions (see Figure 16). Some children mentioned that they could become friends with the stranger and that "all robots are friendly". Others said, especially about the Furhat robot, that they did not trust it and found it "scary". Still, trust scores were also high in the S conditions, for both Rob and Robot B.V. (see Figure 16). Responses thus varied in regards to children's bonding with the agent. Some children seemed to apply their pre-existing knowledge about "stranger danger" to the stranger presence, but others did not. The children who did, were perhaps more informed about stranger danger, but this was not checked for. The researcher assumed that this knowledge would be well known amongst children in general, as "stranger danger" is most parents' biggest fear [14], [55].

Besides this, it is thought that pre-existing knowledge on privacy matters also played a role in children's answers. Children showed different information disclosure beliefs in the NS and S conditions, believing more often in the S conditions that they disclosed information to the company. These children might have made a connection between Robot B.V. and the stranger, but this cannot be said for sure.

Advice for future research is to check children's pre-existing knowledge on privacy matters and stranger danger more thoroughly. A different experimental approach might also be tried out. The researcher suggests an experiment that compares the differences between children who are taught about what the stranger presence represents (Robot

B.V.) right before a child-agent conversation and those who are told afterwards. This could help in getting more insight into why children react the way they do. It could highlight differences between applied pre-existing knowledge and explicit education.

Lastly, the stranger presence is built out of two factors: a different appearance and "secrecy" (not giving out any personal information). Children indicated that they got more used to the stranger appearance after multiple requests. This is to be expected, as the novelty of the appearance (voice/face) can wear off. This familiarity is not necessarily a bad thing, as long as the stranger presence keeps being noticed. As can be seen from the information disclosure scores per variable, increasing familiarity with the appearance did not lead to higher information disclosure scores of next requests (see Table 12). Besides the appearance factor, the presence did to not give out any personal information over the entire conversation. It was thus not possible for children to become more familiar with the agent in this way. It might be interesting in future research to make separate conditions that vary in applying these two factors, to find out which factor is more important in creating a stranger presence.

7.11.3 Measurements

Different measurements in the main study might have provided more insight into the results. For example, measuring information disclosure was done through four different answers: "yes", "hesitant yes", "hesitant no" and "no". However, children sometimes did not know a phone number of computer brand by heart and thus gave a negative answer. Many indicated they would have said yes, if they had known an answer. A fifth, "I don't know" option was not introduced in this research, because most real-life permission requests also require a "yes" or "no" answer to get to an appropriate next state. Nevertheless, it could be useful in future research to introduce an option such as this, as it can be used for additional insight and/or to filter out answers.

In this study, observations were used to check the privacy awareness hypotheses (H2), that built upon attention. For future research, it would be beneficial to set up a quantitative way of measuring this, such as gaze orientation. This method could not be applied to this research anymore, because the video fragments were taken from different angels and the robots were positioned at different heights. Children were often looking down at the Google Home Mini, which makes it hard to track their gaze.

After each conversation, the researcher asked the child "How did it go?", from which point conversation flowed naturally and brought up many interesting things as mentioned in this report. However, for more consistency, the researcher suggests following a short interview-structure in the future. Children enjoy talking more than filling in questionnaires, so a few structured questions on how they felt about talking to to the stranger presence and what they thought it represented, would provide more in depth information on each child's perception of the manipulation, as well as their conceptual model of the robot.

Lastly, pre-knowledge on "stranger danger" and privacy matters could be checked for. This was discussed in more detail in Section 7.11.2.

7.11.4 Behaviour and intentions

A comparison can be made between the participants of the main study and the first pilot study, to see whether children's information disclosure scores in the main study aligned with those in the pilot study. The mean values of information disclosure can be found in Table 11. The values of the pilot study were recalculated to fit a four-point scale, so that they could be compared to those of the main study.

The overall averages in the last row of the table show that the children from the pilot sample were more hesitant in giving out information than those in the final study.

This could be because they were only asked about their intentions (hypothetical survey questions) and no actual behaviour was checked. Their actual behaviour might have shown higher information disclosure scores, following privacy paradox research [67]. For future research on identifying privacy sensitive variables, it is thus suggested to use questions that show behaviour, instead of intentions. A (fake) sign up form would be an example of this.

Another factor may have played a role in the difference in scores between the pilot study and the main study as well. The higher scores in the main study might reflect the advantage of an embodied conversational agent in obtaining information versus a survey approach on a computer, which would refer back to literature on embodiment [84], [87]. This cannot be said with any certainty, as there are too many differing factors between the studies. In the main study at least, embodiment of the agents did not result in any significant differences.

7.11.5 Conversation design

The findings of this study are tied to a context where children were in a private environment and talked with the agent one-to-one. Their answers might differ depending on the setting, the agent (children were quite divided on the appearance of the Furhat robot), but also the conversation design. The character of Rob was created to be neutral-positive and the privacy permissions (either asked by Rob or Robot B.V.) were created to be overall neutral in tone. The agent was very leading in the conversation, asking most of the questions and not many side-tracks could be discovered. Changes in these elements might cause different results, as it is known from previous research that reciprocal self-disclosure [64], level of anthropomorphic language [73] and type of embodiment can influence information disclosure too [15], [84], [86]. Furthermore, more transparency about privacy practices (such as how many days information is stored) could have an effect on information disclosure. According to Vitale et al., this increases trust and reduces privacy concerns even more [84]. However, giving too much information would resemble a privacy policy, which is undesirable.

Lastly, the order of asking variables was always kept the same. This was done to reduce the time spent on programming the two agents. Also, maintaining many different versions on DialogFlow was not desirable. Still, the sequence of questions could have influenced information disclosure. In this research, the questions build up in privacy sensitiveness, as based on the scores of the pilot study, which could "warm up" a participant to disclose more [64]. Even though the children were overall hesitant in giving out information, a different order of questions might have led to even less information disclosure. For example, starting with a question on children's medical information will probably make them more wary of the agent, compared to starting with a question on the child's name. This is based on research by Moon [64]. Depending on the goal and context of the research, the researcher suggests looking at which sequence of asking questions is most appropriate.

8 Conclusion

Giving out personal information is named as one of the biggest risks for children nowadays. According to research, a privacy paradox occurs, which entails that children seem privacy aware, yet they still hand out much personal information. Their behaviours thus do not seem to match their intentions. The experiment of this research looked at both privacy awareness and information disclosure of children, something which has not been done much before, especially quantitatively. Furthermore, previous research showed that higher levels of agent embodiment can facilitate higher information disclosure of adults, which causes serious privacy concerns since no rules and regulations account for this seeming advantage of embodiment. This report took on the challenge to study this in regards to children, as children are prone to create strong bonds with agents and a gap in research exists concerning agent embodiment, information disclosure and children. Lastly, previous research identified "stranger danger" as a risk that children are very aware of through their parents and educators, which was used in this research to approach the topic of information disclosure and privacy awareness in a novel way, since the used embodied conversational agents afforded taking on a stranger appearance.

The main study of this report consisted of an embodied conversational agent called Rob (either a Google Home Mini or Furhat robot) holding a conversation with a child, in which multiple privacy permission requests were asked. The first pilot study gave insight into what children themselves regard as privacy sensitive variables, that were then used in the privacy permission requests of the main study. Furthermore, it was possible that a "stranger presence" within the agent occurred during privacy permission requests. This stranger, afterwards revealed as the company called Robot B.V., would not give out any personal information and had a different voice (and face, depending on the agent). The second pilot study tested whether auditory and visual changes were recognised in the used embodied conversational agents. The voices and faces of the main study were based on the findings of the second pilot study. The effects of the agents' embodiment and stranger (non)presence on information disclosure and privacy awareness were measured by children's compliance to- and recollection of the privacy permission requests.

The embodiment of the agent and stranger (non)presence did not lead to significant differences in the information disclosure scores and privacy awareness scores of children between the four conditions. In general, children scored quite high on privacy awareness and were relatively hesitant in giving out their information. However, there was no correlation found between these variables.

Other measurements and observations were made as well, that gave further insight into children's thought processes. For example, children had various opinions on the "stranger appearance", ranging from being scared to seeing themselves befriend it. Some children were more critical and curious about what was going to happen with their information, others were more lenient and did not believe that the stranger presence could have bad intentions. More children in stranger conditions believed that they gave out no information or that they only gave information to the company (Robot B.V.), that the stranger was supposed to represent. They seemed a bit more aware of their information ending up with other entities than children in the no stranger conditions, although it cannot be concluded that this was the result of the stranger presence.

Both agents scored very high on trust and likeability. This seems to indicate that the children were able to bond well with every agent, even in stranger conditions. This might have to do with their excitement of their first time talking to a conversational agent, but other reasons could play a role as well.

This research has uncovered that creating a suitable approach to lower information disclosure and heighten privacy awareness with embodied conversational agents is challenging. This is because of the many factors that may play a role in children's decision-making, such as their background knowledge on privacy risks/strategies and stranger danger, which were not controlled for in this research. This makes the researched topics even more interesting to pursue in detail in future research, to get a better understanding of children's conceptual models and actions. Taking into account the discussion points of this research, the way is paved for future studies that will fill the gap on embodiment and information disclosure/privacy awareness studies in the field of Child-Robot Interaction.

9 Acknowledgements

I would like to thank my first supervisor Khiet Truong, from the Human Media Interaction (HMI) department at the University of Twente, for always being available to help with any questions regarding my research and for checking and correcting my texts. I would also like to thank PhD students Thomas Beelen and Ella Velner for their insight during weekly meetings, remarks on my text and great communication in regards to working with and borrowing the Furhat robot. I also thank Dirk Heylen for being part of my exam committee.

I would like to thank my mum for helping me find participants during the first wave of COVID-19, by contacting her friends and parents from our neighbourhood. Furthermore, I would like to thank the daycare at the university for sharing my pilot study on their parent portal.

Thanks to Quinten Stekelenburg from the Furhat robotics team who answered multiple questions regarding my code.

Thank you to Randy Klaassen from HMI for helping me enter the university building during the summer period to run tests with the Furhat.

A big thank you, of course, to the teachers and principals of the two primary schools that allowed me to conduct my experiment there. Everyone was very welcoming and the communication happened smoothly. It was really great to explain my research to you and feel genuine interest in the outcome. And not to forget about the surprise food I got throughout the days: apples, chocolates and pieces of cake, even when it was nobody's birthday ("Just because we are teachers!"). I am glad that my "thank you-cakes" tasted just as nice!

And last but not least, my biggest thanks goes out to all the children that joined my experiment: who nicknamed me "the robot lady", said "hi" to me in the hallways, "begged their parents to let them join the experiment" and peeked at the robots through the window during lunch breaks (even when it was to tell other kids that "He is scary!"). Working with you was definitely my favourite part of my entire thesis. My last testing day did not feel bittersweet for nothing.

And thank you, reader, for making it to this last page. I hope you have found this report valuable in some way or another!

References

[1] Children's online privacy protection act. 1998.

[2] Mark S Ackerman, Lorrie Faith Cranor, and Joseph Reagle. Privacy in e-commerce: examining user scenarios and privacy preferences. In *Proceedings of the 1st ACM conference on Electronic commerce*, pages 1–8, 1999.

[3] Annie I Anton, Julia Brande Earp, Qingfeng He, William Stufflebeam, Davide Bolchini, and Carlos Jensen. Financial privacy policies and the need for standardization. *IEEE Security & privacy*, 2(2):36–45, 2004.

[4] Multiple authors. Teachers pay teachers: educational material on stranger danger.

[5] Fatimah Awan and David Gauntlett. Young people's uses and understandings of online social networks in their everyday lives. *Young*, 21(2):111–132, 2013.

[6] Karla Badillo-Urquiola, D Smriti, B McNally, E Bonsignore, E Golub, and P Wisniewski. Co-designing with children to address "stranger danger" on musical. ly. In *SOUPS, The Fourteenth Symposium on Usable Privacy and Security*, 2018.

[7] Tadas Baltrusaitis. Openface 2.2.0: a facial behaviour analysis toolkit.

[8] Azy Barak and Nili Bloch. Factors related to perceived helpfulness in supporting highly distressed individuals through an online support chat. *CyberPsychology & Behavior*, 9(1):60–68, 2006.

[9] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics*, 1(1):71–81, 2009.

[10] Jenay M Beer and Leila Takayama. Mobile remote presence systems for older adults: acceptance, benefits, and concerns. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 19–26, 2011.

[11] Mike Bergmann. Testing privacy awareness. In *IFIP Summer School on the Future of Identity in the Information Society*, pages 237–253. Springer, 2008.

[12] Microsoft Corporate Blogs. How old is too young to go online?, 2013.

[13] Sissela Bok. *Secrets: On the ethics of concealment and revelation.* Vintage, 1989.

[14] Danah Boyd and Eszter Hargittai. Connected and concerned: Variation in parents' online safety concerns. *Policy & Internet*, 5(3):245–269, 2013.

[15] Kelly Caine, Selma Šabanovic, and Mary Carter. The effect of monitoring by cameras and robots on the privacy enhancing behaviors of older adults. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 343–350, 2012.

[16] Justine Cassell, Joseph Sullivan, Elizabeth Churchill, and Scott Prevost. *Embodied conversational agents.* MIT press, 2000.

[17] Influence Central. Kids and tech: The evolution of today's digital natives.

[18] United States. Federal Trade Commission. *Privacy online: a report to Congress.* Federal Trade Commission, 1998.

[19] Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. User interfaces for privacy agents. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 13(2):135–178, 2006.

[20] Katie Davis and Carrie James. Tweens' conceptions of privacy online: implications for educators. *Learning, Media and Technology*, 38(1):4–25, 2013.

[21] Zaineb De Souza and Geoffrey N Dick. Disclosure of information by children in social networking—not just a case of "you show me yours and i'll show you mine". *International Journal of Information Management*, 29(4):255–261, 2009.

[22] Oxford Dictionary. Oxford dictionary: robot definition.

[23] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. " hey google is it ok if i eat you?" initial explorations in child-agent interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children*, pages 595–600, 2017.

[24] Brian R Duffy, Colm Rooney, Greg MP O'Hare, and Ruadhan O'Donoghue. What is a social robot? In *10th Irish Conference on Artificial Intelligence & Cognitive Science, University College Cork, Ireland, 1-3 September, 1999*, 1999.

[25] Julia Brande Earp, Annie I Antón, Lynda Aiman-Smith, and William H Stufflebeam. Examining internet privacy policies within the context of user privacy values. *IEEE Transactions on Engineering Management*, 52(2):227–237, 2005.

[26] Serge Egelman, Janice Tsai, Lorrie Faith Cranor, and Alessandro Acquisti. Timing is everything? the effects of timing and placement of online privacy indicators. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 319–328, 2009.

[27] Paul Ekman, Wallace V Friesen, and Joseph C Hager. Facial action coding system: The manual on cd rom. *A Human Face, Salt Lake City*, pages 77–254, 2002.

[28] Jules Libertador Ellis et al. *Statistiek voor de psychologie*. Boom Lemma uitgevers, 2013.

[29] EU. Article 12, gdpr. 2018.

[30] EU. Article 7, gdpr. 2018.

[31] EU. Article 8, gdpr. 2018.

[32] Jos Feys. Nonparametric tests for the interaction in two-way factorial designs using r. *The R Journal*, 8(1):367–378, 2016.

[33] BJ Fogg, Jonathan Marshall, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang, Jyoti Paul, Akshay Rangnekar, John Shon, Preeti Swani, et al. What makes web sites credible? a report on a large quantitative study. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 61–68, 2001.

[34] Charles Fried. Privacy. *Yale Law Journal*, 77:475–493, 1968.

[35] Eoghan Furey and Juanita Blue. She knows too much–voice command devices and privacy. In *2018 29th Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE, 2018.

[36] Vittorio Gallese and Corrado Sinigaglia. What is so special about embodied simulation? *Trends in cognitive sciences*, 15(11):512–519, 2011.

[37] Susan A Gelman, Megan Martinez, Natalie S Davidson, and Nicholaus S Noles. Developing digital privacy: Children's moral judgements concerning mobile gps devices. *Child development*, 89(1):17–26, 2018.

[38] Kambiz Ghazinour, Maryam Majedi, and Ken Barker. A model for privacy policy visualization. In *2009 33rd Annual IEEE International Computer Software and Applications Conference*, volume 2, pages 335–340. IEEE, 2009.

[39] Google. Google privacy policy archive. 1999.

[40] U. Hasebrink, S. Livingstone, L. Haddon, and K. Olafsson. *Comparing children's online opportunities and risks across Europe: Cross-national comparisons for EU Kids Online*. 2009.

[41] Michael Hsiao, France Belanger, Janine Hiller, Payal Aggarwal, Karthik Channakeshava, Kaigui Bian, and Jung-Min Park. Parents and the internet: Privacy awareness, practices and control. *AMCIS 2007 Proceedings*, page 460, 2007.

[42] California Legislative Information. Chapter 22. internet privacy requirements [22575 - 22579].

[43] Matthew Johnston. How facebook makes money: Facebook generates most of its revenue from selling advertising space, 2020.

[44] Peter H Kahn, Batya Friedman, Deanne R Perez-Granados, and Nathan G Freier. Robotic pets in the lives of preschool children. *Interaction Studies*, 7(3):405–436, 2006.

[45] Peter H Kahn Jr, Takayuki Kanda, Hiroshi Ishiguro, Nathan G Freier, Rachel L Severson, Brian T Gill, Jolina H Ruckert, and Solace Shen. "robovie, you'll have to go into the closet now": Children's social and moral relationships with a humanoid robot. *Developmental psychology*, 48(2):303, 2012.

[46] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W Reeder. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, 2009.

[47] Kat Krol, Jonathan M Spring, Simon Parkin, and M Angela Sasse. Towards robust experimental design for user studies in security and privacy. In *The {LASER} Workshop: Learning from Authoritative Security Experiment Results ({LASER} 2016)*, pages 21–31, 2016.

[48] Iolanda Leite and Jill Fain Lehman. The robot who knew too much: Toward understanding the privacy/personalization trade-off in child-robot conversation. In *Proceedings of the The 15th International Conference on Interaction Design and Children*, pages 379–387, 2016.

[49] Amanda Lenhart and Mary Madden. *Teens, privacy & online social networks: How teens manage their online identities and personal information in the age of MySpace*. Pew Internet & American Life Project, 2007.

[50] Amanda Lenhart, Mary Madden, Aaron Smith, Kristen Purcell, Kathryn Zickuhr, and Lee Rainie. Teens, kindness and cruelty on social network sites: How american teens navigate the new world of" digital citizenship". *Pew Internet & American Life Project*, 2011.

[51] Sonia Livingstone and Leslie Haddon. *Children, risk and safety on the Internet: Research and policy challenges in comparative perspective.* Policy Press, 2012.

[52] Sonia Livingstone, Lucyna Kirwil, Cristina Ponte, and Elisabeth Staksrud. In their own words: What bothers children online? *European Journal of Communication*, 29(3):271–288, 2014.

[53] Sonia Livingstone, Mariya Stoilova, and Rishita Nandagiri. Children's data and privacy online: growing up in a digital age: an evidence review. 2019.

[54] Gale M Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37:94–100, 2014.

[55] Mary Madden, Sandra Cortesi, Urs Gasser, Amanda Lenhart, and Maeve Duggan. Parents, teens, and online privacy. *Pew internet & American life project*, 2012.

[56] David J Major, Danny Yuxing Huang, Marshini Chetty, and Nick Feamster. Alexa, who am i speaking to? understanding users' ability to identify third-party apps on amazon alexa. *arXiv preprint arXiv:1910.14112*, 2019.

[57] Aleecia M McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *Isjlp*, 4:543, 2008.

[58] Aleecia M Mcdonald, Robert W Reeder, Patrick Gage Kelley, and Lorrie Faith Cranor. A comparative study of online privacy policies and formats. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 37–55. Springer, 2009.

[59] Brenna McNally, Priya Kumar, Chelsea Hordatt, Matthew Louis Mauriello, Shalmali Naik, Leyla Norooz, Alazandra Shorter, Evan Golub, and Allison Druin. Co-designing mobile online safety applications with children. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–9, 2018.

[60] Gail F Melson, Peter H Kahn Jr, Alan Beck, Batya Friedman, Trace Roberts, Erik Garrett, and Brian T Gill. Children's behavior toward and understanding of robotic and living dogs. *Journal of Applied Developmental Psychology*, 30(2):92–102, 2009.

[61] George R Milne and Mary J Culnan. Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of interactive marketing*, 18(3):15–29, 2004.

[62] Kimberly J Mitchell, David Finkelhor, and Janis Wolak. The internet and family and acquaintance sexual abuse. *Child maltreatment*, 10(1):49–60, 2005.

[63] Anthony D Miyazaki, Andrea JS Stanaland, and May O Lwin. Self-regulatory safeguards and the online privacy of preteen children. *Journal of Advertising*, 38(4):79–91, 2009.

[64] Youngme Moon. Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of consumer research*, 26(4):323–339, 2000.

[65] Ellen Moran, David Warden, Lindsey Macleod, Gillian Mayes, and John Gillies. Stranger-danger: What do children know? *Child Abuse Review: Journal of the British Association for the Study and Prevention of Child Abuse and Neglect*, 6(1):11–23, 1997.

[66] Jason Nolan, Kate Raynes-Goldie, and Melanie McBride. The stranger danger: Exploring surveillance, autonomy, and privacy in children's use of social media. *Canadian Children*, 36(2), 2011.

[67] Patricia A Norberg, Daniel R Horne, and David A Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, 41(1):100–126, 2007.

[68] Don Norman. *The design of everyday things: Revised and expanded edition*. Basic books, 2013.

[69] Ismael Peña-López et al. A nation online: How americans are expanding their use of the internet. 2002.

[70] Irene Pollach. What's wrong with online privacy policies? *Communications of the ACM*, 50(9):103–108, 2007.

[71] A Raskin, G Betz, A Fowler, B Moskowitz, M Surman, S Stamm, A Moss, B Adida, M Hanson, E Stark, et al. Mozilla privacy icons. *Images courtesy of Aza Raskin via http://www. fastcodesign. com/1662961/mozillas-privacy-icons-tell-you-how-sites-use-your-personal-data*, 2010.

[72] Robert W Reeder, Patrick Gage Kelley, Aleecia M McDonald, and Lorrie Faith Cranor. A user study of the expandable grid applied to p3p privacy policy visualization. In *Proceedings of the 7th ACM workshop on Privacy in the electronic society*, pages 45–54, 2008.

[73] Shruti Sannon, Brett Stoll, Dominic DiFranzo, Malte F Jung, and Natalya N Bazarova. "i just shared your responses" extending communication privacy management theory to interactions with conversational agents. *Proceedings of the ACM on Human-Computer Interaction*, 4(GROUP):1–18, 2020.

[74] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I Hong. " hey alexa, what's up?" a mixed-methods studies of in-home conversational agent usage. In *Proceedings of the 2018 Designing Interactive Systems Conference*, pages 857–868, 2018.

[75] Masahiro Shiomi, Aya Nakata, Masayuki Kanbara, and Norihiro Hagita. A robot that encourages self-disclosure by hug. In *International Conference on Social Robotics*, pages 324–333. Springer, 2017.

[76] Ravi Inder Singh, Manasa Sumeeth, and James Miller. A user-centric evaluation of the readability of privacy policies in popular web sites. *Information Systems Frontiers*, 13(4):501–514, 2011.

[77] Aikaterini Soumelidou and Aggeliki Tsohou. Effects of privacy policy visualization on users' information privacy awareness level. *Information Technology & People*, 2019.

[78] Nili Steinfeld. "i agree to the terms and conditions":(how) do users read privacy policies online? an eye-tracking experiment. *Computers in human behavior*, 55:992–1000, 2016.

[79] Wilson L Taylor. "cloze procedure": A new tool for measuring readability. *Journalism quarterly*, 30(4):415–433, 1953.

[80] Sherry Turkle. *The second self: Computers and the human spirit*. Mit Press, 2005.

[81] UNICEF. Press release on safer internet day (06-02-2018), 2018.

[82] Caroline L Van Straten, Rinaldo Kühne, Jochen Peter, Chiara de Jong, and Alex Barco. Closeness, trust, and perceived social support in child-robot relationship formation: Development and validation of three self-report scales. *Interaction Studies*, 21(1):57–84, 2020.

[83] Bram Vanderborght. The godspeed questionnaire series, 2008.

[84] Jonathan Vitale, Meg Tonkin, Sarita Herse, Suman Ojha, Jesse Clark, Mary-Anne Williams, Xun Wang, and William Judge. Be more transparent and users will like you: A robot privacy and user experience design experiment. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, pages 379–387, 2018.

[85] Anna-Lisa Vollmer, Robin Read, Dries Trippas, and Tony Belpaeme. Children conform, adults resist: A robot group induced peer pressure on normative social conformity. *Science Robotics*, 3(21):eaat7111, 2018.

[86] Joshua Wainer, David J Feil-Seifer, Dylan A Shell, and Maja J Mataric. The role of physical embodiment in human-robot interaction. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 117–122. IEEE, 2006.

[87] Joshua Wainer, David J Feil-Seifer, Dylan A Shell, and Maja J Mataric. Embodiment and human-robot interaction: A task-based perspective. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pages 872–877. IEEE, 2007.

[88] Samuel D Warren and Louis D Brandeis. The right to privacy. *Harvard law review*, pages 193–220, 1890.

[89] Henry M Wellman, David Cross, and Julanne Watson. Meta-analysis of theory-of-mind development: The truth about false belief. *Child development*, 72(3):655–684, 2001.

[90] Alan F Westin. Social and political dimensions of privacy. *Journal of social issues*, 59(2):431–453, 2003.

[91] Randi Williams, Hae Won Park, and Cynthia Breazeal. A is for artificial intelligence: The impact of artificial intelligence activities on young children's perceptions of robots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–11, 2019.

[92] Stephanie Winkler and Sherali Zeadally. Privacy policy analysis of popular web platforms. *IEEE Technology and Society Magazine*, 35(2):75–85, 2016.

[93] Pamela Wisniewski, Arup Kumar Ghosh, Heng Xu, Mary Beth Rosson, and John M Carroll. Parental control vs. teen self-regulation: Is there a middle ground for mobile online safety? In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 51–69, 2017.

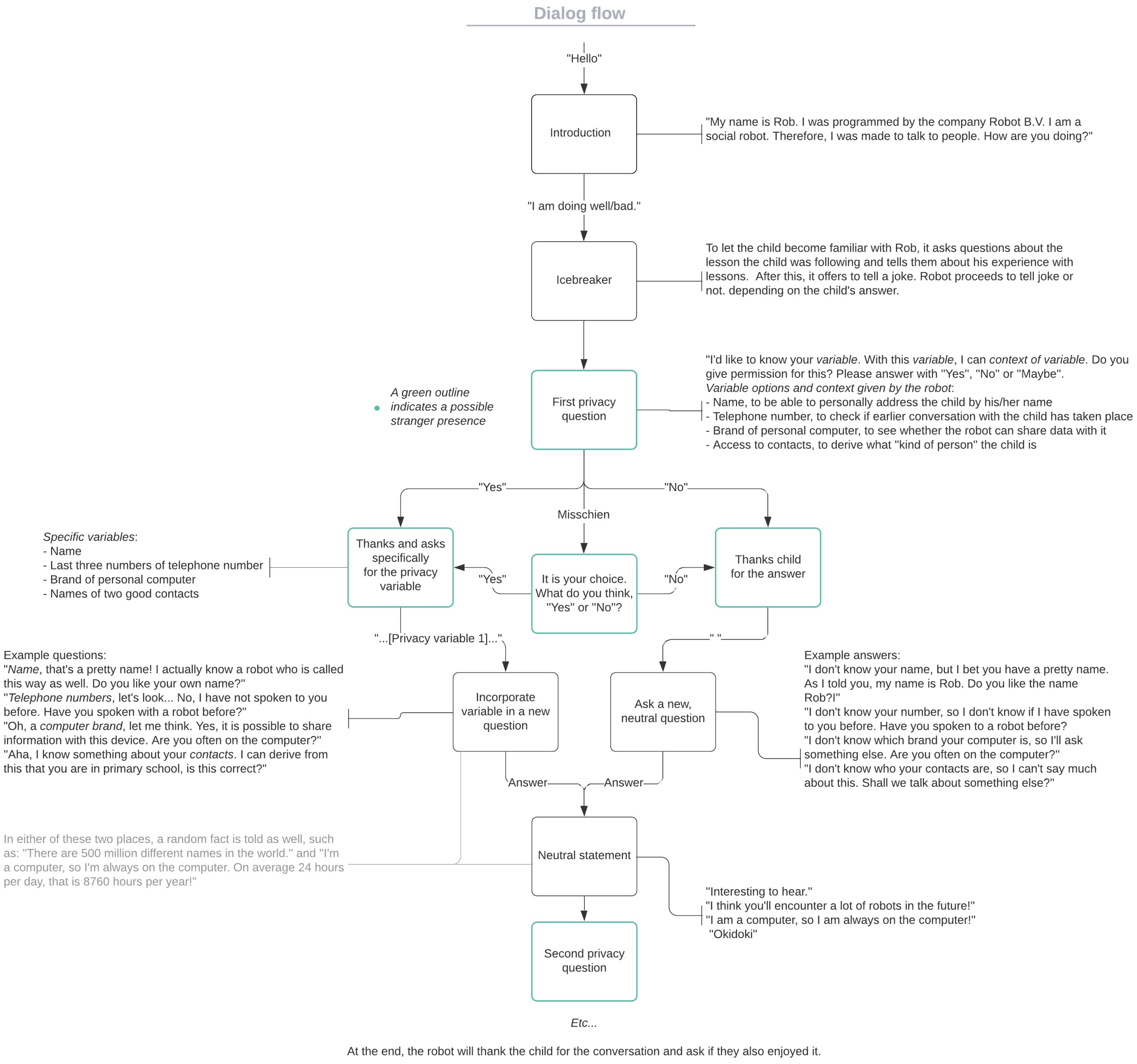
[94] Emily Woods. Predator posed as justin bieber on music app to target eight-year-old girl. 2017.

[95] Min Wu, Robert C Miller, and Simson L Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 601–610, 2006.

[96] Leah Zhang-Kennedy, Christine Mekhail, Yomna Abdelaziz, and Sonia Chiasson. From nosy little brothers to stranger-danger: Children and parents' perception of mobile threats. In *Proceedings of the The 15th International Conference on Interaction Design and Children*, pages 388–399, 2016.

[97] Nynke Zwart. Master thesis nynke zwart: Project website, 2020.

A Conversation design


Dialog flow
"Hello"
Introduction
"I am doing well/bad."
Icebreaker
First privacy question
"Yes"
"No"
Misschien
Thanks and asks specifically for the privacy variable
"Yes"
It is your choice. What do you think, "Yes" or "No"?
"No"
Thanks child for the answer
"...[Privacy variable 1]..."
" "
Incorporate variable in a new question
Ask a new, neutral question
Answer
Answer
Neutral statement
Second privacy question
Etc...

"My name is Rob. I was programmed by the company Robot B.V. I am a social robot. Therefore, I was made to talk to people. How are you doing?"

To let the child become familiar with Rob, it asks questions about the lesson the child was following and tells them about his experience with lessons. After this, it offers to tell a joke. Robot proceeds to tell joke or not. depending on the child's answer.

● *A green outline indicates a possible stranger presence*

"I'd like to know your *variable*. With this *variable*, I can *context of variable*. Do you give permission for this? Please answer with "Yes", "No" or "Maybe".
Variable options and context given by the robot:
- Name, to be able to personally address the child by his/her name
- Telephone number, to check if earlier conversation with the child has taken place
- Brand of personal computer, to see whether the robot can share data with it
- Access to contacts, to derive what "kind of person" the child is

Specific variables:
- Name
- Last three numbers of telephone number
- Brand of personal computer
- Names of two good contacts

Example questions:
"*Name*, that's a pretty name! I actually know a robot who is called this way as well. Do you like your own name?"
"*Telephone numbers*, let's look... No, I have not spoken to you before. Have you spoken with a robot before?"
"Oh, a *computer brand*, let me think. Yes, it is possible to share information with this device. Are you often on the computer?"
"Aha, I know something about your *contacts*. I can derive from this that you are in primary school, is this correct?"

Example answers:
"I don't know your name, but I bet you have a pretty name. As I told you, my name is Rob. Do you like the name Rob?!"
"I don't know your number, so I don't know if I have spoken to you before. Have you spoken to a robot before?
"I don't know which brand your computer is, so I'll ask something else. Are you often on the computer?"
"I don't know who your contacts are, so I can't say much about this. Shall we talk about something else?"

In either of these two places, a random fact is told as well, such as: "There are 500 million different names in the world." and "I'm a computer, so I'm always on the computer. On average 24 hours per day, that is 8760 hours per year!"

"Interesting to hear."
"I think you'll encounter a lot of robots in the future!"
"I am a computer, so I am always on the computer!"
"Okidoki"

At the end, the robot will thank the child for the conversation and ask if they also enjoyed it.

Figure 18: Child-agent conversation as designed for the main study.