Robot-Mediated Interviews: A robotic intermediary for facilitating communication with children

Luke Jai Wood

Submitted to the University of Hertfordshire
in partial fulfilment of the
requirements of the degree of
Doctor of Philosophy
under schedule A

October 2014

Acknowledgments

I gratefully acknowledge the support I received from the University of Hertfordshire in the form of a scholarship.

I would like to thank my supervisors Kerstin Dautenhahn, Austen Rainer, Ben Robins and Hagen Lehmann, who have provided me with excellent guidance, support and encouragement throughout the duration of this PhD. I would also like to thank Dag Sverre Syrdal, Iolanda Iacono, Sophie Mayhew, Chris Miller, Maria Jesus Hervas Garcia, Rachel Marr-Johnson, Nicholas Shipp and Kinga Grof for providing support with various aspects of the research such as coding data and assisting during the experiments at the school.

I am grateful to the staff and children of the schools who participated in this work along with the experts who participated in the user panels. I would also like to offer a special thanks to two specialist professionals from the Metropolitan Police, Ann Stuart and Toran Whybrow, who provided valuable input into this research, and Rob Sharp the educational psychologist that conducted the field trial with the robot.

Lastly I would like to thank my parents for their continuing support throughout the duration of the PhD, without whom it would have been a lot more difficult.

Abstract

Robots have been used in a variety of education, therapy or entertainment contexts. This thesis introduces the novel application of using humanoid robots for Robot-Mediated Interviews (RMIs). In the initial stages of this research it was necessary to first establish as a baseline if children would respond to a robot in an interview setting, therefore the first study compared how children responded to a robot and a human in an interview setting. Following this successful initial investigation, the second study expanded on this research by examining how children would respond to different types and difficulty of questions from a robot compared to a human interviewer. Building on these studies, the third study investigated how a RMI approach would work for children with special needs. Following the positive results from the three studies indicating that a RMI approach may have some potential, three separate user panel sessions were organised with user groups that have expertise in working with children and for whom the system would be potentially useful in their daily work. The panel sessions were designed to gather feedback on the previous studies and outline a set of requirements to make a RMI system feasible for real world users. The feedback and requirements from the user groups were considered and implemented in the system before conducting a final field trial of the system with a potential real world user.

The results of the studies in this research reveal that the children generally interacted with KASPAR in a very similar to how they interacted with a human interviewer regardless of question type or difficulty. The feedback gathered from experts working with children suggested that the three most important and desirable features of a RMI system were: reliability, flexibility and ease of use. The feedback from the experts also indicated that a RMI system would most likely be used with children with special needs. The final field trial with 10 children and a potential real world user illustrated that a RMI system could potentially be used effectively outside of a research context, with all of the children in the trial responding to the robot. Feedback from the educational psychologist testing the system would suggest that a RMI approach could have real world implications if the system were developed further.

Contents

Chapter 1: Introduction

Previous research at the University of Hertfordshire's Adaptive Systems Research Group has focused on robot mediated play for children with autism as part of the AuRoRA project (AUtonomous mobile RObot as a Remedial tool for Autistic children) [1]. Using robots such as KASPAR [2], the group has investigated how robots can instigate collaborative play and social interaction [3, 4]. The research conducted in this dissertation focuses on a novel and largely unexplored application area, establishing how humanoid robots such as KASPAR could be used for Robot-Mediated Communication in an interview setting. To the best of my knowledge this is the first PhD thesis exploring the area of Robot-Mediated Interviews for children. This thesis explored and evaluated how a robotic system such as KASPAR could be used for Robot-Mediated Interviews (RMIs). The children who took part in this research were aged between 7 and 11, both neurotypical children and children with a range of special needs took part in the studies.

In the initial stages of this research it was necessary to establish if children would respond to a robot in an interview setting as this was largely exploratory research, therefore the first study compared how children responded to a robot and a human in an interview setting. Following a successful initial investigation, the second study expanded on this by examining how children would respond to different types and difficulty of questions from a robot compared to a human interviewer. Building on these studies and moving towards the AuRoRA project, the third study investigated how children with special needs would respond to a RMI. The decision to investigate this possibility was heavily influenced by literature indicating that children with special needs are at a higher risk of abuse than neurotypical children [5], although conviction rates are low due to communication difficulties [6]. Since research indicates that children with special needs such as autism can respond positively to robots, investigating this avenue appeared to be a logical step. Following encouraging results from the three studies indicating that a RMI approach may have some potential, three separate user panel sessions were organised with professionals that have expertise in working with children and whom the system would be potentially applicable to. The participants provided feedback on previous studies and outline a set of requirements to make a RMI approach feasible for real world users. A common requirement outlined by each of the three user panels was to increase the flexibility of the system, enabling users to make quick and spontaneous responses to unforeseen avenues of questioning. In an effort to meet this requirement, the system developed for the final study implemented a direct voice conversion process for the interviewer to talk directly through the robot using a microphone. Prior to the final study, the feedback and requirements from the user groups were considered and implemented in the system before conducting a field trial of the system with a potential real world user.

While supervisors provided regular advice on the direction of this research and editorial feedback on written work, the studies conducted in this thesis were designed, conducted, analysed and written by the author of this thesis. Although the term "we" is used in this thesis, it is for stylistic purposes only.

1.1. Motivation

This thesis focuses specifically on investigating how humanoid robots can be used in a formal interview setting with young children. Interviewing children is a relatively infrequent task that is usually carried out for specific reasons. Professionals that occasionally need to interview children include: police, social service workers and educational psychologists. The inspiration for research into the possibility of a RMI approach originates from two primary difficulties experienced by police and social services interviewers which are child comfort and the controllability of the interviewer's facial expressions and body language. When a child is being interviewed by the police or social services it is almost certainly for a good reason, and as a result sometimes the questions that children are asked can be of a sensitive or personal nature, with high levels of stress and anxiety associated to the questions [7],p.57. The first difficulty experienced by interviewers in these interviews is that the child may find it intimidating or scary talking to an adult, which can impede communication. We hypothesise that some children may find it less intimidating talking to a small humanoid robot than an adult, which could have a positive impact on the child's ability or willingness to communicate. The second difficulty relates to the human interviewers ability to control his or her body language and facial expressions. Due to the nature of some interviews, the revelations disclosed can be quite disturbing. While specialist professional interviewers undergo significant amounts of training and development, they are still only human and as a result sometimes have unconscious uncontrollable reactions.

When a child reveals something particularly shocking in an interview it can be very difficult for the interviewer to maintain their composure without subtly indicating their reactions to the child. The Achieving Best Evidence (ABE) document used by UK Police states that displaying signs of shock to a child can have a detrimental effect on the child's disclosure when communicating information about events [7],p.196.

In an entirely different field of study, research exploring the use of social robotics indicates that children often respond well to socially interactive robots [8-10]. From robotic pets and educational aids [9, 11-14] to therapeutic and assistive tools for children [2, 3, 15-19], children generally react well to robots. Since research suggests that children respond well to robots in a variety of settings, this inspired the idea that robots could potentially be used to interview children. Conducting interviews with a robot may assist in solving some of the problems experienced by police and social services interviewers. Because a robot is neither an adult nor an explicit figure of authority, children may find it easier and less intimidating talking to a robot than an adult interviewer. Research by Bainbridge et al. [8] indicates that that robots appear to have more authority than virtual agents, however in this study there is not a comparison between a robot and a human. In relation to the difficulty of unintentional body language experienced by professional interviewers, using a robot could negate this problem because a robot can be fully and consciously controllable, eliminating the possibility of unintentional body language or facial expressions. The overall hypothesis of this research is that children may be more comfortable talking to a robot because it would not be viewed as an authority figure or display unintentional body language and facial expressions, these two positive factors may assist communication, particularly where conventional methods are proving to be unsuccessful.

Another area of socially interactive robotics that shows particular promise is research aimed at assisting children with special needs such as autism to encourage social interaction [3, 9]. Research on the AuRoRA project at the University of Hertfordshire has focused on investigating how robots can be used to assist in play and therapy with children that have autism [1]. Much of this research indicates that children with autism respond well to robots and that robots can be used to assist in breaking the children's isolation [3, 4, 10]. Focusing on RMIs, research indicates that children with a disability are up to four times more likely to be a victim of abuse than typically developing children [5], however the number of cases resulting in prosecution is relatively low [11, 12]. Interviewing children with special needs can be very difficult, particularly when talking about a sensitive or emotionally provocative topic or event. A Mencap review states "Those who can't communicate can't tell, and those who can't communicate well won't be believed" [6]. Given that considerable amounts of research indicate that children with special needs respond well to robots, it was felt that investigating how children with special needs such as autism respond to a robot in an interview setting would also be constructive. Therefore establishing if robots could also be used to assist communication with children that have special needs was one of the objectives of this research.

Shortly after commencing research in this field, a study conducted by Bethel et al. in 2011 [13] investigated if neurotypical children aged 4 and 5 years old, were as likely to share a secret with a NAO robot [14], as they were with an adult. The quantitative results from this study were inconclusive, however the qualitative results revealed that the children would readily interact with the robot and speak to it in a similar manner as they would with an adult. The research in this thesis builds on the work of Bethel et al. and thoroughly investigates how children respond to robots in a formal interview setting.

The research presented in this thesis focuses on assistive technology and aims to determine if robots could be useful tools for interviewers conducting interviews with children. It is thought that approximately 16 per cent of the Western population experience some form of serious maltreatment during their childhood [15]. Communicating with young children about potentially traumatic or emotionally provocative events can sometimes be difficult and as a result gathering information from these witnesses can prove to be very upsetting for the child and sometimes impossible. Since research suggests that children respond very well to robots in a variety of settings, it seems plausible that using a robot in this setting may have some advantages over an adult interviewer, particularly in cases where children have a special need that affects their ability to communicate. If robots can be used to assist child witnesses disclose information, this may potentially help safeguard children in the future and therefore this would appear to be a very worthwhile cause for investigation.

The foreseen goal of this research was not investigating ways of replacing human interviewers, but to investigate the possibility of providing professionals with a robotic tool that could serve as an interface to create an enjoyable and comfortable setting for children to talk to. RMIs, as described in this thesis, could allow the professional to precisely control the robot's behaviour (e.g. facial expressions, body language), which is often very hard to do even for professionally trained interviewers, in particular when the topic of the interview may be emotionally sensitive, see Section 2.1. An interview in this research is defined by the author of this thesis as: a one to one verbal exchange of information between the interviewer and interviewee on a specific topic or subject. The purpose of the interview will be clearly focused on information acquisition without other tasks to distract the interviewee which is where this research differentiates itself from Bethel et al.'s research [16]. The research conducted in this thesis is asserting a proof of concept for the potential of an RMI system rather than testing how effective an RMI would be in a sensitive setting, as it is important to first test if the concept of an RMI would work.

1.2. Research questions

The purpose of the research conducted in this thesis is to establish if robots could be used as intermediaries to facilitate communication with children in an interview setting. To determine this, the following primary research questions will be addressed in this thesis:

RQ-1) To what degree do children respond to a robot in an interview setting?

It is important to first establish if children will respond to questions from a robot in a formal interview setting. As this is exploratory work there are no guarantees that children will respond to questions asked by a robot. For example, the children may choose to ignore the robot and remain silent.

RQ-2) How do children respond to a robot compared to a human interviewer?

Children merely responding to a robot in an interview setting is not enough. If a robot is to be useful in a real world setting it is important to examine how children respond to a robot compared to a human, particularly in terms of the quality of children's verbal responses, as this can sometimes be the basis on which important decisions are made. For example, the children may talk to the robot a lot more than the human interviewer, but the information they disclose might be completely different or possibly fabricated, thus it is important to check the consistency and frequency of the information that the children provide to a robot compared to a human interviewer.

RQ-3) Will question type and difficulty impact on children's responses to a robot in an interview setting?

Since the nature of the foreseen application area has the potential to be sensitive, establishing how children respond to different types and difficulties of questions is imperative. In a legal setting, factual questions are likely to be the most important, whereas in a social services or child welfare setting, establishing how a child feels may also be important. Children only responding to easy questions from a robot would not be enough if a robot were to be useful for real world applications, therefore investigating how children respond to different types and difficulty of questions is necessary.

RQ-4) How do children with special needs respond to a robot compared to a human interviewer?

Children with special needs are more likely to be victims of maltreatment than neurotypical children. Since considerable amounts of research indicate that children with special needs respond well to robots, it would seem logical to investigate how children with special needs respond to a robot that is interviewing them too.

RQ-5) Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?

Because of the exploratory nature of this work, ascertaining if real world users think that a RMI approach could be useful for them in the future is one of the most important questions to answer, as this determines which application domains should be targeted and indicates where research efforts should be focused.

RQ-6) What would experts working with children require from a Robot-Mediated Interviewing system?

Experts working with children quite often are not roboticists or computer scientists. For a RMI system to be useful in the real world, the system would need to be sufficiently flexible and user friendly to be useable by non-technical staff. Therefore, gathering the relevant information as to what real world users would require from such a system is essential for producing a useful system with real world potential.

RQ-7) Could robots such as KASPAR be used in an interview setting effectively by an expert working with children?

To ultimately answer the overall primary research question, could robots be used as intermediaries for facilitating communication with children in an interview setting? The system would need to put in the hands of an expert to gather feedback on how the system performs and whether the expert considers that a RMI approach would be useful in a real world situation.

Table 1.2.1: Studies conducted in this research were designed to answer specific research question. This table defines which research questions each study aimed to answer

Study/Description	RQ
Chapter 4 – How do young children respond to a humanoid robot in an interview setting (Study #1) This study investigated if children would respond to a robot in an interview setting and examined how children's responses towards the humanoid robot differ in comparison to their interaction with a human interviewer.	1, 2
Chapter 5 – Testing the impact of question difficulty on information recovery with a robotic interviewer (Study #2) The second study verified the results of the first study whist building on this research by examining if question type and difficulty would impact information recovery when being interviewed by a robot.	1, 2, 3
Chapter 6 – How do children with special needs respond to a robotic interviewer (Study #3) The third study investigated how children with a variety of special needs would respond to a robot compared to a human interviewer using the same question set as Study #2.	1, 2, 3, 4
Chapter 7 – Gathering expert opinions (Study #4) This study focused on gathering expert opinions to ascertain if a RMI approach could have real world applications based on previous research conducted in this thesis. In addition to this, a comprehensive specification of what real world users would require from a RMI system was also outlined by the experts.	5, 6
Chapter 8 – Field trial with a potential real world user (Study #5) The final study consisted of a field trial with a potential real world user working with the system. The system had been modified according to the requirements outlined by the experts in the previous study. Feedback was gathered from the expert trialling the system about the usability of the system and the potential for real world applications.	1, 2, 5, 6, 7

1.3. Overview of thesis

Chapter 2: This chapter provides the background research and motivations for this work before concluding with the proposed research questions this thesis aims to answer.

- **Section 2.1:** Outlines common practices and best approaches for interviewing both neurotypical children and children with special needs, before highlighting some of the difficulties of interviewing children and iterating what scope there is for aids that can assist the interview process.
- **Section 2.2:** Provides an overview of relevant studies and principles in the field of social robotics in relation to the potential of a RMI approach for young children. Key studies and important factors that need consideration are also addressed in the latter parts of this section.
- **Section 2.3:** Outlines the research questions, defining where this research aims to contribute to knowledge, and hypotheses about the expected outcome of this research, taking into consideration the findings from the literature covered in this section.
- **Section 2.4:** Summarises the research covered in chapter 2 and outlines the motives for conducting the research in the thesis.

Chapter 3: This chapter provides in-depth details of the pre-study methodological considerations for the first three studies in this research that initially investigated and tested the concept of RMIs. The elements covered in this chapter include, details of the experimental setup implemented, an overview of the robot, design of coding schemes, specifics of transcription analysis methods, and details of the questionnaires used. The planning and preparation for these studies were vital to the success of this research because it ensured that there would be an abundance of data to support this research.

Chapter 4: This chapter presents the work and subsequent results from Study #1. Since there were no guarantees that children would respond to a robot in an interview setting, the first study conducted examined how children's responses towards the humanoid robot KASPAR in an interview setting differ in comparison to their interaction with a human in an identical setting. Twenty-two children aged between 7 and 9 took part in the study. Each child participated in two interviews, one with an adult and one with a humanoid robot. Measures include coding of the children's behaviour during the interviews, transcriptions and questionnaire data. The questions in the interviews focused on a special event that had recently taken place in the school. The results of this study contributed to knowledge by establishing that the children would interact with a humanoid robot in an interview setting, furthermore we found that the children interacted with a robot in a very similar manner to how they interacted with a human interviewer. The temporal analysis revealed that the most notable difference between the interviews with KASPAR and the human were the duration of the interviews, the eye gaze directed towards the different interviewers, and the response time of the interviewers.

Chapter 5: This chapter presents the work of Study #2, following up the previous study which indicated that children respond to a robot in a very similar manner to which they do a human interviewer. This follow-up study investigated how twenty-three children (aged between 7 and 9) respond to questions of varying difficulty from a robot compared to a human interviewer. Similar to the first study each child participated in two interviews, one with an adult and one with a robot. The main questions in these interviews focused on the theme of pets and animals. After each interview the children were asked to rate the difficulty of the questions and particular aspects of the experience. Measures include coding of temporal data, transcripts of what the children said and questionnaire data. Results from the quantitative data analysis contribute to knowledge by revealing that the children interacted with KASPAR in a very similar manner to how they interacted with the human investigator, and provided both interviewers with similar information and amounts of information regardless of question difficulty.

Chapter 6: The first two studies in this research investigated how typically developing children responded to a robot in an interview setting. This chapter builds on the first two studies and presents the work of Study #3 which investigated the possibility of RMIs with children that have various special needs. Children with a disability are up to four times more likely to be a victim of abuse than neurotypical children, however the number of cases that result in prosecution is relatively low. One factor influencing this low prosecution rate is the communication difficulties encountered when questioning children with special needs. This study investigated how five children with special needs aged 9 to 11 responded to the robot compared to a human interviewer. The measures used in this study included temporal analysis of responses, detailed analysis of transcribed data, questionnaire responses and data from engagement coding. Similar to Study #2, the main interview questions varied in difficulty and focused on the theme of animals and pets. The results from quantitative data analysis reveal that some of the children appeared to have a preference towards KASPAR, however all of the children provided both interviewers with similar information regardless of question difficulty. Further to this the qualitative analysis suggests that some children may have been more engaged with the robot. The contribution to knowledge from this study is that some children with special needs, in particular communication difficulties, appear to respond better to a robot than a human interviewer in terms of engagement. The design and implementation of an alternative means of analysis that was designed to capture how engaged the children were during the interaction is also outlined in this chapter, along with a discussion of the potential advantages and pitfalls of such an approach.

Chapter 7: Following three successful studies, this chapter outlines the results of Study #4, a set of three expert user panels aimed at establishing what real world users working with children would require from a RMI system. The user panels were conducted with several potential user groups to gather views of the current system and ascertain what they would require for the system to be useful to them. The user panels consisted of specialist child protection police officers, intermediaries who specialise in working with very young children or children with communication difficulties, health care professionals and legal experts. In order to test if a RMI approach would work in a real world setting, it is important to determine what the experts working with children would require from such a system. The findings of these user panels contribute to knowledge by establishing that professionals from a variety of backgrounds need a flexible system that allows for any path of questioning to take place, and that the system should have a small degree of autonomy to allow the user to focus on interviewing the child rather the controlling the robot. In addition to this, the areas and children for which this type of system may be most applicable were also outlined.

Chapter 8: The first three studies investigated if and how children world respond to a robot in an interview setting. The fourth study investigated what potential real world users would require from a RMI system. This chapter presents the work of Study #5 where KASPAR was placed in the hands of an educational psychologist for a field trial. The primary aim of this concluding study was to gather feedback on the final system that had been developed and ultimately establish if this is an approach that could be useful for real world users. Before the trial took place some pre-trial questions were asked to gather information about the psychologist's role and ideas about the potential advantages and pitfalls of a RMI approach. The field trial took place with ten child participants aged between 7 and 8. Although the results of the user panels suggested that a RMI system would most likely be useful for children with special needs, the purpose of this field study was to test the usability of the system and establish if a RMI system could be used by a potential real world user, therefore it was not essential to work with children with special needs. The questions were focused on a 3 minute cartoon that the children watched prior to the interview taking place. Upon completion of the interviews, the educational psychologist was asked some post-trial questions to establish how he felt using the system and ultimately, his opinion on whether this type of system could have real world potential. The feedback and contribution to knowledge from the educational psychologist about the system was that a RMI approach may have the potential to be used in a real world setting by a number of different professionals, provided all technical issues were fully resolved and the system was very robust.

Chapter 9: The last chapter summarises the studies and research in this thesis, before drawing to a conclusion and outlining the future potential of RMIs.

1.4. Contribution to knowledge

Interviewing children can be a challenging task for professionals, particularly when the questions that are being asked are of a sensitive or emotionally provocative nature. The research outlined in this thesis aims to investigate if robots such as KASPAR could be useful tools for conducting interviews with children. In conducting this research the following contributions have been made:

Human-Robot Interaction

The research conducted in this thesis has shown that children will respond to a humanoid robot in a formal interview setting. In all four of the studies conducted at schools, the children always responded to the robot, with the quantitative data suggesting that the children responded to the robot in a similar manner as they would a human interviewer. Further to this, we have found that question type and difficulty did not appear to impact on the information recovered from the children. These contributions to knowledge would likely be of interest to HRI researchers and psychologists working with children because very little research has investigated how robots can be utilised in an interview setting.

For robotics research to have real world impact it is important that HRI aspects are investigated from all users' perspectives. Another HRI contribution to knowledge from this research investigated the needs of the professionals who would potentially have a use for a RMI system, in order to establish what would be required from an RMI system to work in a real world setting as opposed to a research context. We found that the most important features of a RMI system to potential users would be flexibility, stability and usability. This contribution is likely to be of interest to HRI researchers who are interested in developing RMI platforms.

Assistive robotics

Whilst reviewing literature of techniques for interviewing children, it quickly became apparent that communicating with children that have special needs can be particularly challenging and that children with special needs are more likely to be victims of maltreatment often because of the communication difficulties they incur [6]. Since considerable amounts of research indicate that children with special needs appear to respond well to robots [3, 9], the possibility of using a RMI approach was also investigated. The contribution to knowledge for assistive robotics in this thesis comes from Study #3 conducted with children that have special needs. We found that some of the children appeared to be more engaged with the robot than the human interviewer, highlighting that this approach may have some potential. The findings from this study coupled with previous research of assistive robotics indicate that a RMI approach may be worth further investigation and may yield some positive results. This contribution to knowledge is likely to be of interest to many groups including HRI researchers, psychologists and professionals working with children that have special needs.

Child interview communication aids

Overall this research has investigated the possibility of using a new communication aid in interviews with children. The 2011 Achieving Best Evidence document used by the UK Police force, lists: drawings, pictures, photographs, symbols, dolls, figures and props, as potentially useful interview aids [7],p.89. The research in this thesis has investigated how humanoid robots such as KASPAR could potentially be utilised as an interview aid by experts to conduct interviews with children. Although all of the studies with children in this research were conducted in schools about a non-sensitive topic, the feedback from user panels and field trial with a potential real world user all indicate that a RMI approach may have some potential. Before a RMI system is used in a real world setting, it is important to first test the concept in a safe environment such as a school, which was the purpose of this research. This contribution to knowledge is likely to be of interest to individuals who conducts interviews with young children such as educational psychologists, social service workers and possibly police officers.

Chapter 2: Background and motivations

This chapter reviews literature around common practices for interviewing both neurotypical children and children with special needs, whilst highlighting some of the difficulties experienced when conducting interviews. The current applications of HRI are explored in relation to the potential advantages a robot could possess in an interview setting to assist professional interviewers. Although interviewing children and HRI are two very different disciplines, a considerable amount of research indicates that children often respond well to robots in a variety of settings. Since children appear to respond well to robots, it would seem plausible that children may also respond well to robots in an interview setting. The starting point for investigating the possibility of RMIs was that children may be more comfortable talking to a robot than an adult interviewer, because the child may view the robot as a peer [17] and consequently feel more at ease with the robot.

2.1. Interviewing children

2.1.1. Current approach

Interviewing children can be a challenging task and is therefore carried out by highly trained experts. When a child is involved in a formal interview it is generally for a very good reason, with emotionally provocative questions often needing to be asked, which can be quite stressful for the child. When conducting interviews or training police officers to conduct interviews, the UK police often refer to a document called "Achieving Best Evidence in Criminal Proceedings" (ABE) [7, 18]. The ABE was drafted for the UK's Home Office by a team of experts from varying backgrounds including psychology, law and social services. The guidelines outlined in the ABE have been well researched and are recognised as providing an effective structured method for interviewing people, including young children. Therefore, when structuring and setting the interview questions for studies in this research, the relevant guidelines of the ABE were followed as closely as possible. For example, the information we were gathering would not be used in court, therefore assessing if the children knew the difference between the truth and a lie was not essential and was not established in the studies conducted as part of this research as we did not want to cause the children any undue stress. The ABE suggests that interviews should have four phases: rapport building, free narrative recall, questioning, and closure. A structural diagram of this process can be seen in Figure 2.1.1.1.

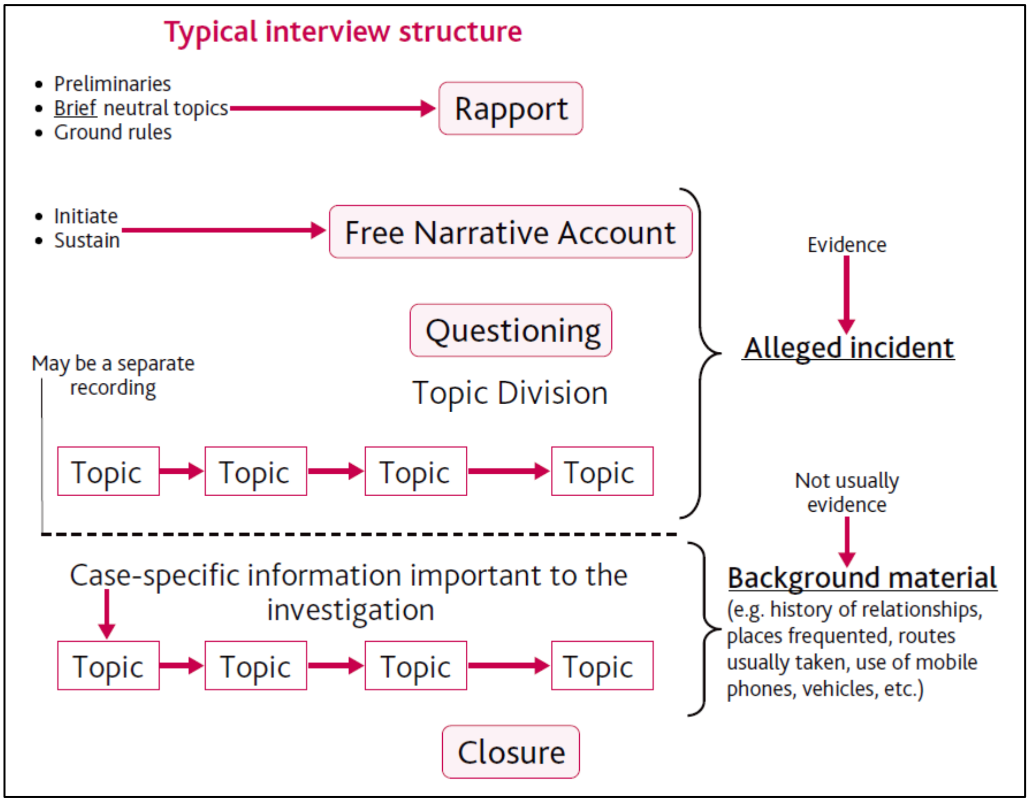


Figure 2.1.1.1: The typical 2011 Achieving Best Evidence Interview Structure [7],p.69.

2.1.1.1. Rapport building

The rapport building phase allows the child to become acquainted with the interviewer. Acquiring a good rapport with the child often assists information recovery in the interviews. The ABE suggests that neutral topics should be discussed in the rapport building phase to relax the child. Spratling et al. [19] reviewed 18 studies and states that "Establishing a connection with the child during the interview was critical to successful research outcomes", reaffirming the importance of establishing a good rapport with the child. A study exploring the experiences of school age children with a Tracheostomy [20] found that allowing children to talk about matters that are important to them, rather than focusing on what is important to the researcher aided in establishing rapport and often produced rich data. The rapport building phase should also be used to set the ground rules for the rest of the interview. The ground rules of a police interview are designed to emphasise that the interviewer was not at the event that is being discussed and does not know all of the details, thus it is important to explain the incident as clearly and in as much detail as possible. Part of setting the ground rules also includes emphasising the importance of telling the truth. In police interviews it is unacceptable for a witness to tell lies, it is therefore explained to the child that they must only say things which are true during the interview. To ensure that a child knows the difference between the truth and a lie, the child is given some examples and asked to answer whether the example given is the truth or a lie. This is because it is vitally important to assess the credibility and reliability of a witness during criminal proceedings. An additional consideration when the police are working with a child witness is the child's perception. Often when a child is interviewed by the police they immediately assume that it is because they have done something wrong. It is important for this to be addressed at an early stage so that the child knows that they are not in trouble.

When conducting the research in this thesis it was beneficial to consider some aspects of the ABE such as using neutral topic questions to build a rapport with the child. Other aspects such as ensuring the child knows they have done nothing wrong would need to be modified for the studies in this research. For example, the interviews in the studies conducted as part of this research were not performed by a police officer therefore the child is unlikely to worry about having done something wrong. However, the child may be concerned that they are being graded or assessed, therefore reassuring the children that they would not be graded was vital. With regards to establishing the children's understanding of truth and lies, this would not be such an important aspect for these studies. It may in fact be beneficial not to state that the child must tell the truth when conducting studies to assess the reliability of their unprompted responses. If a child chooses to lie to the robot but not to the human, this would be an early indication that a robot may not be suitable as an interviewing tool. However, if the child reveals the same information to a robot that they do a human interviewer, this is a good indication that a robot may be a suitable tool for interviewing children.

2.1.1.2. Free narrative recall

The free narrative recall phase is designed for the child to recall as much information as possible without prompting, using minimal direction. The ABE explains that information gathered from free narrative recall is thought to be the most accurate, and would be considered more reliable as evidence in a courtroom. However, the ABE also states that for less articulate witnesses it may be necessary to employ other techniques, provided the technique employed are within the legal framework.

The free narrative recall would be an important consideration when conducting comparative studies. Although the information gathered in these studies would not be relied upon in court, it would be useful to allow the children to express themselves freely, as this could be used as a partial measure to establish how much information children freely provide to a robot compared to a human.

2.1.1.3. Questioning

In police interviews a questioning phase follows after the free narrative recall, which is designed to build on the information gathered in the free narrative recall and help to create a more complete picture. The interviewer focuses on trying to recover key pieces of information that the child may have overlooked in their recall of the event. This allows the interviewer to maximise the amount of useful information they can recover from the child. The questions used in the interview must not be in any way leading, as this would compromise the integrity and legitimacy of the statement from the child. When conducting studies it would be desirable to follow the same procedure to ensure that information recovery is genuine rather than lead.

2.1.1.4. Closure

The ABE suggests that towards the end of any interview there should be a closing phase where the topic of conversation should move back towards a neutral topic, similar to the rapport phase, before thanking the child for their time and concluding the interview.

2.1.2. Additional interview techniques

There are a number of techniques that can be used in interviews, some of which are explored by Geiselman et al. [21]. Geiselman et al. investigated and compared the effectiveness of three interview methods with 89 adult participants: a standard police interview, a cognitive interview and an interview using hypnosis. The participants observed a video clip of a violent crime and were then interviewed by a police officer using one of the three methods 48 hours after the video was shown to them. The measurements used in this study were the number of facts accurately recalled and the number of errors made during recollection. In contrast to a typical police interview, during a cognitive interview the interviewer tries to mentally reinstate the environmental and personal context of the incident for the witness, possibly by asking them about their feelings and general activities on the day. In a cognitive interview the witness is asked to report every detail, even if they consider it to be insignificant because such details might act as a trigger for key information about the incident. Further to this, witnesses are asked to describe the incident from different perspectives, explaining what they think other witnesses may have seen. Geiselman et al.'s study found that the average number of correctly recalled facts for the cognitive interview was the highest with 41.2, followed by hypnosis with 38.0 and 29.4 for the standard interview. There were no significant differences found in the number of errors in each condition. This study suggests that the cognitive interview leads to better memory recall for events, with witnesses being able to recall more relevant information compared with a traditional interview method. The cognitive interview technique is described in the ABE [7],p.92, however it is a specialist technique that requires specific training and is only used in particular circumstances [7],p.92, therefore the most widely used approach was adopted for the purposes of this research.

2.1.3. A chat application compared to a face to face interview

An alternative approach to conducting an interview is via a chat application on a PC. Oskarsson et al. [22] conducted a study in which 64 participants watched a 53 second video clip of a crime taking place and were then interviewed about the video in either a face to face interview, or via a PC chat application. Oskarsson et al.'s study aimed to examine if an alternative interview technique would reduce the errors in witnesses' memory recovery, compared to a traditional face to face interview. The study examined the number of correct details outlined by the participants, the misinformation effects on participants and the perceived confidence of the participants in their statements. The results from this study indicated that the communication via the chat application was as effective as questioning face to face, with both interview techniques yielding equivalent results. A difference was found between male and female participants, showing that female participants use more words in both interview techniques than the male participants. This study relates to the research conducted in this thesis however, the participants in Oskarsson et al.'s study were adults and the alternative method of interviewing using a PC chat application is likely to be cumbersome for younger children. The use of a chat application may however be more useful with older children, particularly teenagers, where texting and instant messaging is often a common form of communication.

2.1.4. Inherent difficulties of interviewing children

When police officers are conducting interviews with young children that have been through a stressful or traumatic ordeal it can be difficult for the interviewer to maintain their composure. Despite their extensive training, it is sometimes inevitable that the interviewer may subtly and unintentionally indicate their thoughts and feelings to the child through body language or facial expressions. The information that a child reveals in an interview can sometimes be quite shocking or surprising. The ABE states "the interviewer should not display surprise at information as this could be taken as a sign that the information is incorrect" [7],p.196. Maintaining such emotional discipline can be quite difficult for a human interviewer, but would be easy for a humanoid robot whose expressions are explicitly controlled.

It is also important that an interviewer does not appear to assume that someone is guilty "So far as possible, the interview should be conducted in a 'neutral' atmosphere, with the interviewer taking care not to assume, or appear to assume, the guilt of an individual whose alleged conduct may be the subject of the interview" [7],p.66. Furthermore, recent research suggests that body language can play a role in misleading witnesses. A study by Gurney et al. [23] found that participants who received positive nonverbal feedback whilst being interviewed were more confident with their answers than participants that received negative nonverbal feedback. In this study positive nonverbal feedback was a subtle nod of the head and negative feedback was a subtle shake of the head.

Nonverbal behaviours such as facial expressions and hand gestures are often produced automatically and spontaneously [24-26]. Gurney et al. [23] concludes that "common nonverbal behaviours (head nodding and shaking) that are likely to occur in interviews can have an impact on eyewitnesses' confidence judgements", and highlights that "By altering the confidence witnesses attribute to their testimony, police investigators can manipulate precisely the quality that eyewitnesses are often judged upon". In a courtroom, jurors often place a lot of trust in confident eyewitness [27], therefore it is important to ensure that both the questions and body language of the interviewer are not leading. Using a robot to interview a person could eliminate any of the subtle unintentional signs in body language that a human interviewer may give away, as the body language of the robot can be fully and precisely controlled by the interviewer.

In addition to body language the mere perception of a person's authority can sometimes have an effect on a witness, particularly with regards to suggestibility [7],p.56. The ABE suggests that reducing the perceived authority differential between interviewer and witness may reduce the possibility of a witness complying with a leading question, "Paying attention to the appropriate form of address at this initial greeting phase can help send a message of equality both now and throughout the interview. This is essential as it reduces the perceived authority differential between interviewer and witness, so that witnesses are less likely to comply with leading questions. As no interview can be perfect, it is essential to build resistance against inappropriate questions, which may unwittingly be used by an interviewer later in the interview" [7],p.187. Using a robot could negate this problem because the robot is clearly not an adult and may not be viewed in the same way.

2.1.5. Interviewing children with special needs

Jones et al. [5] conducted a systematic review of 17 papers and concluded that children with a disability are up to four times more likely to be a victim of abuse than children without disabilities, however the number of cases that result in prosecution is relatively low [11, 12]. Interviewing children with special needs can be very difficult, particularly when talking about a sensitive or emotionally provocative topic or event. A Mencap review states that "Those who can't communicate can't tell, and those who can't communicate well won't be believed" [6]. The ABE suggests that when interviewing children with special needs or communication difficulties, it is advisable to seek advice from a specialist who is familiar with the specific procedures for working with children that have a disability or communication difficulty [7],p.172. Because children with a disability can often have difficulty communicating, sometimes props and intermediaries are used to help facilitate communication. Props suggested by the ABE include: drawings, pictures, photographs, symbols, dolls and figures [7],p.89. Props may be used of a number of reasons:

- To aid in assessing a child's language or understanding
- To keep a child calm and settled
- To support the recall of events
- To enable a child to give an account of events

However, using props must be approached with caution as there are risks and pitfalls associated with the use of props [7],p.89. The risks associated with using props include: potential legal challenges, distortions or inaccuracies (mostly associated to dolls), the potential to stimulate play or fantasy (associated to teddies or animals), and the risk of upset to the carer or child from explicit use of dolls or drawings. Nevertheless, when used appropriately props can be useful tools for interviewing when common techniques are proving ineffective.

2.2. Potential impact of increased accessibility of technology

Techniques and procedures for interviewing child witnesses have improved over the years, but there is still considerable room for improvement. The increased availability and accessibility of social robots means that using robotic technology in interviews with children is now a real possibility. Shibata summarises how robots are being used in a social context, and highlights some of the design issues along with the impact of cross-cultural preferences in the subjective evaluation of socially interactive robots [28]. Shibata explains that social robots are assessed subjectively on their ability to create "high value in interactions with humans". According to Shibata there are four main categories of social robots designed for psychological enrichment which are, "performance robots", "teleoperated performance robots", "operation, building, programming, and control robots", and "interactive autonomous robots". In the case of an interview setting, a teleoperated performance robot could be used as an interface (or prop) to aid information recovery when interviewing a child.

2.2.1. Child-Robot Interaction with social robots

Studies investigating Human-Robot Interaction have steadily increased, from robotic pets [29-31], to educational aids in schools [32, 33], to therapeutic and assistive tools for children [3, 34, 35]. Scheeff et al. [36] conducted a study investigating how age and gender affected peoples interactions with a small mobile robot called Sparky. The study was conducted in both a lab environment and an open environment, with no prior instruction being given to the participants. Sparky was teleoperated from a remote location and could express some emotions such as happiness, fear and anger. The study found that the user group who displayed the most interest in the robot were children, with young children showing particularly high levels of engagement with the robot. Older boys were often more aggressive towards Sparky, and would only respond if the robot showed aggression back towards them, girls on the contrary would pet the robot. The findings of this work suggest that children, particularly young children, are often willing and will actively approach robots to interact with them, indicating that children may find robots engaging interaction partners in a variety of settings.

Building on this research, Tanaka et al. [17] investigated how toddlers socialises with a QRIO robot in a preschool setting. The study was conducted over a five month period with toddlers aged between 18 and 24 months, each child experienced 45 sessions with the robot, which lasted approximately 50 minutes per session. The rational for working with this age group was that the children were unlikely to have preconceptions about robots, and socialisation at such a young age is not as dependant on speech. The study consisted of three phases. In the first phase conducted over 27 sessions, the robot could select behaviours from its full repertoire of behaviours. In the second phase conducted over 15 sessions, the robot was programed to exhibit interesting but highly predictable behaviours. In the Final phase conduced over 3 sessions, the robot was programed to use its full repertoire of behaviours again. The findings from this study revealed that the children engaged with the QRIO robot and after some time would treat the robot as a peer rather than another toy. This finding was verified by a control condition with 2 other toys, one of which resembled the QRIO robot. Furthermore, despite the QRIO robot being the least huggable toy, it was found that the children would hug the robot more than the other 2 control toys. The finding that younger children would interact with the robot as a peer is of particular interest. It was found that children's tactile behaviours with the robot converged towards the type of tactile interactions they would have with their peers. One of the possible reasons for this was that the teachers encouraged the children to be careful with the robot and this may have had an influence over their behaviour towards the robot. Taking this into consideration for the research in this thesis, it will be important to ensure that the children are encouraged to view the robot as a peer rather than another toy, as viewing the robot as a toy may encourage the children to slip into a fantasy type scenario, "Some props, e.g.

teddies, animals, dolls houses, may engender play or fantasy" [7],p.89. Encouraging the children to view the robot as a peer may assist in ensuring that the information that the children provide to the robot is as reliable as possible.

Other studies have focused on investigating how robots can be used to help educate children. Kanda et al. [32] conducted an 18-day field trial at a Japanese elementary school with first-grade and sixth-grade children, to investigate the possibility of using robots as social partners, with the intention of teaching the children English. Two Robovie robots were located in the school for the children to interact with during recess. The robots could gesture with their limbs and articulate more than 300 sentences in English to communicate with the children. The robot could also recognize approximately 50 English words. To measure how successful the robot had been in teaching the children they were given three English tests, a pre-test, to assess the children's starting level of knowledge, a mid-test (1 week into the trial), and a post-test after the trial. From an observational perspective it was found that many of the children quickly lost interest in the robots and as a result, interaction with the robots fell sharply despite the initial interest. The interactions with the robots after the first week did not affect the children's English scores, however the children that maintained their interactions with the robots did see an improvement in their English test scores. It was found that children, who already had an interest in English and knew some English before interacting with the robots, would be more inclined to maintain their interactions with the robots. It is thought that the robots' initial impact created unrealistically high expectations for the children, and as a result many of the children tried to interact with the robots at the same time, which overwhelmed the robots as they were not designed cope with so many users simultaneously. The majority of the children would try to interact with the robots accompanied by one or more friends. The research from these studies provides indications that children of varying ages are happy to interact with and talk to robots in a variety of settings and contexts. Consequently these research papers hint that children may also respond well to robots in an interview context.

2.2.2. Evaluating Child-Robot Interaction in contrast to Child-Adult

Various studies and research articles suggest that children seem to enjoy interacting with robots [36-38]. Nevertheless it is also important to establish how comfortable children are speaking to a robot compared to a human, as this is an integral part of ascertaining if robots could be useful tools for interviewing young children. Recent work has explored how children interacted with iCat robots [39-41]. Shahid et al. [42] investigated if children perceive playing with a robot to be more like playing with a friend, than playing with a toy. In this research the difference between children playing a card game by themselves, with a friend or with an iCat robot was examined. The study used two independent measurements for evaluating the children's interactions. The first measure was a subjective fun rating from the children just after their session had finished. The second method of evaluation employed 70 adults to participate in two perception experiments, where they evaluated if the children had just won or lost a game from a set of clips. The subjective fun scores suggested that children enjoy playing with the robot more than playing alone, but not as much as when playing with a friend. The perception tests indicated that children were more expressive when playing with the robot than when playing alone, but less expressive than when they played with a friend.

This work relates to the proposed area of research in this thesis in terms of assessing how children view robots. The desirable outcome in the research conducted by Shahid et al. would have been for the children to view the iCat robot as a friend rather than a toy. However, the findings suggested that children enjoy playing with a friend more than a robot, and were more expressive with a friend than a robot, but did enjoy playing with the robot more than playing alone. Factors influencing this result included the age and familiarity of the children's friends. The children had known their friends long before the iCat robot, and were of a similar age to their friend. Both of these factors may have contributed to children having more fun with their friends than the robot. Nevertheless, the children enjoyed the company of the robot, and this in itself is a positive sign when establishing if children are likely to be comfortable with talking to a robot.

A further example illustrating how children will interact and talk to a robot can be seen in a case study conducted by Nishio et al. [43]. The study investigated how a teleoperated android could be used to represent a personal presence and how familiarity can affect the interactions. The robot used HI-1, was classified as a geminoid by its creators due to its design being focused on representing an accurate clone of a real person (Ishiguro). The study was conducted with two children who had very different levels of familiarity with Ishiguro. One child was a boy aged 4 who was not familiar with Ishiguro, the other child was a girl aged 10, who was the daughter of Ishiguro. The experiment contained two conditions, one condition was having the real person, Ishiguro, interact with the children, whilst the other condition was to have the geminoid interact with the children. Both conditions were tested twice, so there were 4 sessions per child, 8 sessions in total. The children engaged in conversational tasks on preselected topics with either the geminoid or Ishiguro. After each session, the children had a post interaction interview with one of their parents, to uncover their thoughts on the interactions with Ishiguro and the geminoid. Both of the children were scared of the robot at first, however they both gradually adapted to interacting with the robot over the course of time. The children also had very different views of the geminoid, the older girl quickly realised it was a robot controlled by her father. The younger boy initially insisted that the real Ishiguro was a robot but it was alive and breathing, but later the boy said he was mistaken. The younger boy was very confused, and he described the robot as very serious and said it listened well to his story.

The findings of this research illustrate that a robot can perform conversational tasks with young children relatively well. However, to maximise the potential advantages that a robot may have, it is important to ensure that the child is comfortable with the robot. As stated in the ABE, the importance of building a good rapport with the interaction partner is critical to the success of an interview [7],p.70. Evidently from the work conducted by Nishio et al. the appearance of the robot plays a pivotal role in helping to establish a good rapport, therefore the selection of an appropriate robot to perform conversational tasks with the children is paramount.

Given the application area of RMIs, it would be useful to ascertain how children perceive robots, as this is likely to affect how useful robots will be for information acquisition. Fior et al. [38] investigated if children could form relationships with robots as friends by interviewing children after they had been interacting with a modified robotic arm robot. 184 children aged 5 to 16 individually interacted with a robotic arm robot that had been customised to give the appearance that it had a face. The robot had been pre-programed and was controlled by a researcher to perform a block stacking routine. After the children's interaction with the robot, they were asked a series of questions relating to their social views of the robot in a short interview. 85.9% of the children thought the robot could be their friend, 67.4% of the children would talk to the robot, and most interestingly 45.7% would share a secret with the robot. The physical activities with the robot in this study are not directly related to the proposed research in this thesis, however the views and attitudes of the children towards the robot are very relevant. The study suggests that children are happy to talk to and view robots as friends, with approximately half of them stating that they would be willing to share a secret with a robot. The statistics from this study would very likely change if the robot and some of the conditions were different. Regardless, these statistics are encouraging and would support the hypothesis that children might be willing to communicate and share information with robots.

2.2.3. Robots working with children that have special needs

Children develop at different rates and the way in which they communicate will vary from child to child, but this variation in communicative abilities fluctuates even more amongst children with special needs such as autism. Autism is a developmental condition that is becoming increasing prevalent, this may be due to a number of factors such as, increased awareness of the condition, and changes in the diagnosis criteria [44, 45]. Autism appears in many different forms and varies in its degrees of severity. It exists as a spectrum of conditions ranging from severe low functioning autism, to high functioning and Asperger's syndrome [46]. One of the most common manifestations of autism and Asperger's is an impaired ability for social communication and interaction [47]. Projects such as Aurora [1] and IROMEC [48] are investigating how robots can be used as social mediators to help facilitate and promote social communication and interaction amongst children with autism. Although autism is just one condition, the spectrum of special needs associated with the condition, particularly the impaired ability for social communication and interaction means that research and studies that have been carried out with children that have autism are likely to be relevant to other children that have conditions which affect their social communication and interaction.

Research in robotic therapy for children with autism seems to suggest that some children do respond well to various kinds of robots, from doll like robots to mobile robots [49]. One such example can be found in a study conducted by Robins et al. [9]. Over several months Robins et al. studied how four children interacted with the robotic doll Robota, with the aim of helping to improve social interaction skills and encourage imitation. There were three phases in the study that progressively became more complex, the first of which was primarily an introduction phase. The second phase consisted of the carer actively encouraging interaction and showing that the robot could imitate their limb movement. In the final phase, the children were given no further encouragement to interact with the robot. The quantitative results varied, however some of the children began to use to the robot as an object of shared attention. In addition to this, after the children had become accustom to the robot, all of the children began to include the researcher, actively sharing their experiences. This paper illustrates that some children with autism respond well to humanoid robots, but also indicates that sometimes children with special needs such as autism need more time to become comfortable with a robot.

In a later study, Robins et al. [3] demonstrated how the humanoid robot KASPAR had been used in long-term studies and one off trials, to support children with autism, focusing particularly on three cases. The first case presents a one off trial in Germany with a six-year-old girl that has severe autism. After the child overcomes her initial hesitation, she began to explore the robot, particularly its face. Further into the trial she also reached out to the researcher, which was not typical of her usual behaviour, as she would usually seclude herself and rarely showed any interest for social interaction. The second case study presented the findings of a long-term trial over several weeks with a young boy who has severe autism. The boy would interact with his family at home, but would not initiate interactions with other people, be it class mates or teachers. In the early stages the child began by exploring KASPAR on a solitary basis, but after several weeks of interacting with the robot he actively sought to involve the teacher in the interaction, sharing his excitement. The third case study examined how KASPAR's long-term interactions with a 16-year-old boy seemed to affect his willingness to interact with others in a positive social manner. Typically the boy was often unwilling to tolerate other people and found it difficult to focus on one activity for a sustained period. However, the boy maintained interactions with KASPAR for prolonged periods and after several weeks, began to engage in imitation games with his therapist. This paper illustrates that some children with autism are willing to interact with humanoid robots such as KASPAR, and that robots such as KASPAR can be used to encourage social interaction. Considering that children with autism often find it difficult to communicate with people, and studies have found that some of these children do respond well to robots such as KASPAR, it is worth investigating how children with autism respond to KASPAR interviewing them.

Surveying research more closely related to interview tasks for children with special needs, a study conducted by Huskens et al. [10] explored how Applied Behaviour Analysis based intervention (ABA) conducted with a robot compared to the same intervention conducted by a human. The purpose of the intervention was to encourage the children to self-initiate question asking, because it is considered that self-initiated behaviours can have beneficial effects on the child’s level of expressive, receptive language and may increase social communication abilities. The results of this study indicated the intervention was successful in both experimental groups with the number of self-initiated questions increasing from the baseline. However no definitive conclusions could be drawn from the data on whether the intervention with the robot was any different to the intervention with the human. The fact that the results of the intervention with the robot were similar to the intervention with a human is encouraging because aspects of this study have similarities to the work in this thesis and would suggest that there may be some potential worth investigating because the children did respond to the robot.

2.2.4. Robots performing information acquisition tasks

In recent years research investigating how robots can be used for information acquisition has begun. Kiesler et al. [50], used a Nursebot to investigate if people would anthropomorphize with a robot more than a projection of a remote robot, or a software robot-like agent present on a screen or projected. Each participant was presented with one of these four conditions in which, the robot or robot-like agent conducted interviews with the participants asking them about their health. The study found that people were more engaged with the robot, but disclosed less information about undesirable behaviour, they also chose healthier food options when offered after the interviews. This research suggests that people anthropomorphized more with the robot than the software agents, regardless of the non-humanoid look of the robot. Anthropomorphization towards a robot in a RMI setting would be desirable, as this may encourage the interviewee to view the robot as a peer. In relation to the research in this thesis, the study was limited as all four of the conditions used a robot or a software agent. Therefore, there is no comparison to show how the information people would provide to a robot would differ to what they would provide to a real person. The investigation by Kiesler et al. also found that people were less likely to reveal information about undesirable behaviour to the robot than the software agent. The research conducted by Kiesler et al. only looked at adult's response to a robot or software agent. The research in this thesis will focus on children, therefore how children view, communicate and interact with a robot is likely to differ greatly from how adults interact with robots. The information that children would share with a robot is also likely to differ from what adults would share with a robot.

In very closely related work Bethel et al. [13, 51] investigated if children aged between 4 and 6 years old, are as likely to share a secret with a humanoid robot as they are with an adult. The research was conducted over two studies that followed a similar procedure on both occasions with the main difference being the robot used. The first pilot study used the ZENO robot which has a realistic human face, while the NAO robot with a more toy like appearance was used in the second study. In the pre-interview phase, a teacher introduced the child to the lead investigator. The investigator asked a number of questions, and clarified the procedure with the child. In addition to this, the child was questioned to ascertain if they knew what a secret was, if the child did not understand what a secret was the term secret was explained to them. The investigator then shared a secret with the child, and explicitly asked the child not to tell anyone. In the second phase of the experiment, the child took part in an interaction task with the robot and another adult separately. During the interactions the child was encouraged to divulge the secret that they had been told to the interaction partner. In the final phase, a post-interaction interview was conducted with the child where the child answered further questions.

The results from these studies revealed that the children would readily interact with the robots and speak to them in a similar manner as they would an adult. The results of the pilot study with the ZENO robot indicated that the children preferred to interact with the adult rather than the robot. However, in the follow up study with the NAO robot, the children preferred to play with the robot rather than the adult. Although the children preferred to interact with the NAO robot, they had a tendency to share the secret with the ZENO robot with the same or less prompting than what they would with the NAO robot. This study has some parallels to the proposed area of research in this thesis. However in the study by Bethel et al. the robot was acting as a social interaction partner rather than explicitly leading an interview. It is unlikely that the children knew that the purpose of their interaction was to gather information from them, as they were participating in a physically interactive task. In a RMI setting, it will be clear to the children that we are trying to gather information from them, as there will not be any other task for them to focus on whilst having the interview.

A later paper by Bethel et al. [16] investigated if an eyewitness testimony would be influenced by a robotic interviewer compared to a human interviewer when presented with post event misinformation. The study consisted of 101 adult participants with an approximate equal distribution into 2 groups, and 2 conditions within each group. One group were interviewed by the NAO robot, whilst the other group were interviewed by an adult, the 2 conditions were controlled vs misinformation. The participants viewed a slideshow of a crime being committed and were then asked some questions about the slide show after. In the misinformation condition the participants were asked questions with misinformation embedded within the questions i.e. “He stopped at his tool box, lifted a hammer, and laid the calculator underneath it”. In the control condition this would have been worded “When he stopped at his tool box and put the calculator inside it, did he appear to be hiding it?”. The results of this study found that the participants were likely to be misled by the human interviewer, but not the robotic interviewer. Although this study was not conducted with children, the results appear to be encouraging for the potential of RMIs, as the study would seem to suggest that a robot could acquire more accurate information from the interviewee, which would be advantageous in a legal setting.

2.2.5. Child-robot interactions in sensitive contexts

Although research about the possibility of RMIs is very limited, other areas of social robotics have investigated the possibility of using robots for interactions in sensitive contexts [34, 52]. In 2010 the ALIZ-E project began with the intention to develop artificial intelligence (AI) for small social robots and investigate how young people respond to and interact with these robots for longer periods of time [53]. The application domain that the project focused on was HRI with diabetic children, investigating how robots can help by offering training and entertainment. Baroni et al. conducted an investigation with parents, children and healthcare workers on what a robotic companion could do for a diabetic child [54]. The results from a series of brain-storming sessions conducted with 22 families that have a diabetic child suggested that a companion robot could fulfil a number of specific functions to support diabetic children in their daily lives. The roles and areas in which it was believed that a robot could assist were: entertainment, self-management, sensitive listening, self-confidence and motivation. The area especially relevant to this research is sensitive listening. It is particularly important to enable diabetic children to share their feelings and emotions. The children and parents believed that a robot could fulfil the role of a sensitive listener, particularly if the robot managed to become familiar enough with the child to be viewed as a peer. The healthcare professionals reviewing and validating the results of the brain-storming sessions were also in agreement with this view. One possible means of implementing this that was detailed in the paper was to integrate "sensitive listener" behaviours during other kinds of interaction such as when the robot is playing a quiz game with the child. The research in this paper illustrates that there is enthusiasm towards the possibility of robots fulfilling information acquisition tasks with children and highlights that this would not necessarily need to be performed in a formal manner, depending on the context and information being disclosed.

2.2.6. Using the right robot for the task

For the proposed area of research it is important to select an appropriate robot for the task. As research by Kiesler et al. suggests, it may be advantageous to have a robot that the interviewee will anthropomorphise with [50]. The most typical conversation partner for a human is another human, therefore it is likely that a humanoid robot would be the most suitable interview partner. There are many different humanoid robots that are used for social robotics, examples include: Asimo [55], Robovie [56], Maggie [57], iCub [58], HRP3 [59], KASPAR [2], HRP-4C [60] and the NAO Robot [14]. An additional important consideration when selecting a robot is how a child will view the robot. If a robot is too big it may be more intimidating to the child and will possibly view the robot as an adult or an authority figure, hence losing the potential advantage of portraying a peer like appearance. Further to this consideration, it is also important to use a robot with a human-like appearance, not just a humanoid form. This is important because a non-human like appearance many encourage the children to slip into a fantasy type scenario. The ABE states that "Some props, e.g. teddies, animals, dolls houses, may engender play or fantasy" [7],p.89. By using a robot with a more human-like appearance it may reduce the possibility of the children slipping into a fantasy type scenario.

Considering these factors and research conducted in social robotics to date, the robot considered most appropriate for this research was the KASPAR robot [2]. Dautenhahn et al. [2] discusses the design details and rationale of the KASPAR robot, along with some examples of where and how KASPAR has been used [3, 61, 62]. Rather than focusing on developing an ultra-realistic android type robot [43], or orienting towards the technologically complex aspects of robotics [63, 64], the aim of KASPAR was to build a robot that could be used safely for Human-Robot Interaction studies. It is explained in the paper that over 400 children in many different studies have met KASPAR, and both neurotypical children and children with special needs have responded to KASPAR positively. The literature and evidence presented provides a good indication as to KASPAR's suitability for the proposed task, with regards to both its technical capabilities and encouraging track record of working with a variety of people, including children with autism.

2.2.7. HRI factors worth considering when conducting studies

When working in the area of Human-Robot Interaction it is important to consider various aspects of the robots behaviour. Sabanovic et al. [65] discusses issues relating to the evaluation of social interactions between humans and robots. Sabanovic et al. reviewed two observational studies involving robots interacting socially with humans. The first robot was a mobile conference-attending robot called GRACE, which was designed to search for a person wearing a pink hat at a conference, and would ask people for directions in order to locate this person. The second robot was a receptionist robot called Roboceptionist, which provided information to visitors through story telling. The comparison of how likely people are to interact with both these robots found that the mobile robot GRACE was more likely to be approached and interfaced with, but this may have been due to its' location and novelty value. The results from this paper highlighted the importance of timely responses to users, particularly with the Roboceptionist. Users will not react to just any movement or response from a robot, the movements and gestures must be conducted in a timely manner. Although both of these robots are very different to KASPAR, as are their intended applications, this is a universally important point to consider regardless of the users age. When conducting research the timely response of KASPAR will be vitally important to facilitating, engaging and maintaining interactions.

In addition to the importance of timely responses, it is also necessary to consider the impact of gestures. An exploratory study by Kose-Bagci et al. [61] worked with 12 adult participants to investigate how a drumming activity with the humanoid robot KASPAR could be used as a tool for social engagement. The purpose of these experiments was not to assess KASPAR's drumming ability, rather it was to assess the impact of non-verbal gestures within the interaction. In this experiment, KASPAR had two modes: listening, in which KASPAR recorded and analysed the rhythms played, and playing, in which KASPAR played the rhythm back using the drum. The effects of KASPAR's social gestures in this turn-taking game were studied to gauge the impact of non-verbal gestures. Three levels of gesturing were used in the study, in the first condition KASPAR imitated the drumming without any gesturing. In the second condition, KASPAR made some simple gestures, but this was dependent upon the user's participation. In the third condition KASPAR used gesturing regardless of the user's participation, KASPAR performed the same drumming task in all conditions. Two methods of assessment were used to analyse the interactions, the first was a qualitative analysis of the participant's experience. The second was a quantitative analysis of the drumming performance. It was found that the most popular game, from the user's perspective was the second condition game with conditional gesturing, and the least favourite was with the condition with no gestures. Closer inspection of the results revealed that there were more drumming errors in the conditions the participants liked the most. An additional factor that influenced the interactions was the gender of the participants, with female participants preferring more gestures than males. Although this study was not conducted with children, it illustrates how factors from both the robot and the user can influence the interactions with the robot and can directly impact on the enjoyment of the interaction. A fundamental lesson that can be taken from this study is the importance of adapting to the user, what is good for one user may not be good for another. It is likely that children with autism will have different interaction preferences to neurotypical children. Therefore, it will be important to consider these preferences when working with children that have autism.

In a later study by Kose-Bagci et al. [66], the effects of embodiment as well as gestures in a drumming game with KASPAR were investigated. This study was conducted on a much larger scale, working with 66 children aged 9 to 10 years. The children were divided into two groups in terms of gestural conditions (with and without gestures). Each child played a drumming game with KASPAR in three different embodiment conditions (physical embodiment, virtual embodiment and disembodiment), with or without gesturing depending upon which condition they had been assigned to. The results from this study suggested that the children enjoyed interacting with a physical robot more than a virtual or disembodied robot. These results are consistent with the study by Kiesler et al. [50] where it was also found that people engaged with a physical robot more than a virtual robot or agent. From a gestural perspective, in the study conducted by Kose-Bagci et al. it was found that gestures played a particularly positive role in the physical embodied condition. This study provides further evidence that children do respond well to the KASPAR robot, and highlights the potential benefits of gestures for enhancing social interactions. However, it is also important to ensure that gestures used in interviews are not leading as this could mislead the interviewee [23, 67].

2.3. Research questions generated from literature

Reviewing the literature regarding child interviews and HRI it is clear that a RMI approach has potential. There appear to be number of research questions that have been raised from the literature that would need to be investigated in order to establish if robots could be useful tools for experts working with children to conduct interviews.

RQ-1) To what degree do children respond to a robot in an interview setting?

The first question to answer is will children even respond to questions from a robot in a formal interview setting and take the activity seriously? Although this would appear trivial, there are no guarantees that children will respond to questions from a robot. HRI research clearly indicates that children are often willing to accept robots and keen to interact with them, however this is usually in a play type context which is very different to a formal interview. The most closely related work by Bethel et al. has investigated how children will share secrets with a robot which is encouraging [51], however, the studies carried out by Bethel et al. were not solely an interview, instead the robot was acting as a social interaction partner rather than explicitly leading an interview. It is unlikely that the children knew purpose of their interaction was to gather information from them, as they were participating in a physically interactive task. Therefore it was essential that we test a robotic interviewer in a setting where it will be clear to the children that we are trying to gather information from them and without participating in any other task whilst having the interview.

RQ-2) How do children respond to a robot compared to a human interviewer?

Since conducting interviews with children is all about gathering information, it is important to establish that using a robot is at least as good as a human interviewer, otherwise there would be no point and the potential advantages with regards to child comfort and controllability of body language would pale into insignificance. Because RMIs is a relatively new field of research, we obviously needed to answer this question too. The work by Bethel et al. is again the most closely related and her papers indicated that people are less likely to be misled by a robot than a human interviewer [16], which again is a potentially positive attribute of a robotic interviewer. If a robot is to genuinely be useful in a real world setting it is important to see how children respond to a robot compared to a human, particularly in terms of the quality of children’s verbal responses, as this can sometimes be the basis on which important decisions are made, particularly when the police are involved. The children may talk to the robot a lot more than the human interviewer, but the information they disclose may be completely different or possibly fabricated, thus it is important to check the consistency and frequency of the information the children provide.

RQ-3) Will question type and difficulty impact on children’s responses to a robot in an interview setting?

As can be seen from the literature regarding interviewing children, the nature of the foreseen application area is inherently sensitive, therefore establishing how children respond to different types and difficulties of question is imperative. Children only responding to easy questions from a robot would not be enough if a robot were to be useful for real world applications. In addition to this, establishing the different types of questions that children will respond to from a robot will also be useful in ascertaining which applications RMIs with children would be most useful for i.e. social service applications or police interviews.

RQ-4) How do children with special needs respond to a robot compared to a human interviewer?

During the literature review we found that children with special needs are often much more likely to be victims of maltreatment than neurotypical children. Since considerable amounts of research indicate that children with special needs respond well to robots, it would seem logical to investigate how children with special needs respond to a robot that is interviewing them too.

RQ-5) Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?

As can be seen from many of the research articles cited in this chapter, gathering users opinions on robotic systems is possibly one of the most important steps if a system is to work in a real world setting. Ascertaining if potential real world users think that a RMI system could be useful for them will assist in determining which application domains should be targeted for this research in the future.

RQ-6) What would experts working with children require from a Robot-Mediated Interviewing system?

Experts working with children quite often are not roboticists or computer scientists. For a RMI system to be useful in the real world, the system would need to be sufficiently flexible and user friendly in order to be useable by non-technical users. The HRI literature reviewed in this chapter illustrates that gathering the relevant information as to what real world users would require from such a system is essential for producing a useful system with real world potential.

RQ-7) Could robots such as KASPAR be used in an interview setting effectively by an expert working with children?

Many of the studies in the literature review focused on researchers using their systems with participants. However, to answer the overall primary research question, "Could robots be used as intermediaries for facilitating communication with children in an interview setting?", putting a RMI system in the hands of an expert working with children to gather feedback on how the system performs is essential. Establishing whether an expert working with children considers that a RMI approach would be useful in a real world situation would be the most effective means of ascertaining the answer for the overall primary research question.

2.4. Summary of literature

The papers reviewed in this chapter have focused on the user group of children, including children with special needs, and an application area where the use of social robots could be beneficial but has been largely unexplored to date. The UK Home Office ABE guidance document for interviews suggests that props may be used in an interview setting where necessary to help facilitate communication [7],p.89. Props could include tools such as humanoid robots, although this approach has not been tried to date. Research in social robotics would suggest that robots may be suitable for tasks such as interviewing children [13], and may have some advantages over a human interviewer, particularly in cases with children that have special needs such as autism [3]. However, before robots are used in highly sensitive cases, it is important to first test the concept of an RMI in a safe environment such as a school, with non-invasive questions. Overall the research indicates that robots could be a useful tool for interviewing children, however this will be assessed by investigating the seven research questions outlined in Section 1.2. Therefore, investigating and utilising current technology to explore the potential benefits it may bring to assist in safe guarding children would seem a worthy cause.

Chapter 3: Generic experimental design and implementation

This chapter discusses the methodology that was used for the first three studies in this research which were designed to investigate and test the concept of an RMI approach. The three studies conducted aimed to answer the first 4 of 7 research questions proposed in this thesis. The elements covered in this chapter include, an overview of the robot, details of the experimental setup implemented, design of coding schemes, specifics of transcription analysis methods and details of the questionnaires used. The planning and preparation for these studies was vital to the success of this research, because it ensured that there would be an abundance of data to support the research and answer the research questions outlined in this thesis. Due to the novelty of the research regarding RMIs for children, the studies conducted in this thesis were exploratory to a large extent.

3.1. Methodological considerations

3.1.1. Preparations

3.1.1.1. Recruiting specialist knowledge

Before conducting the initial studies, two specialist police officers from the Metropolitan Police were consulted with for advice on how to conduct interviews with children. The officers advised on interview structure, phrasing of questions, potential unforeseen circumstances, and directed us towards relevant literature such as the Achieving Best Evidence (ABE) document [7, 18] which police officers in the UK refer to when conducting interviews. The advice imparted by the police officers provided a sound structured foundation for conducting the studies in this research.

3.1.1.2. Procedural design

Careful consideration was given to the structure of the studies to ensure that the data gathered was as reliable as possible. In the first three studies each child experienced two interviews, one with KASPAR and one with the principal investigator, the same interview structure was followed on both occasions. A two-phase counterbalancing method was implemented to reduce the possibility of interview order adversely influencing the results. Half of the children were interviewed by KASPAR first and half were interviewed by the principle investigator first. Counterbalancing was also applied in terms of gender and year group, so that each group had the same number of boys/girls and year groups. Assignment to each of the two groups was otherwise random.

3.1.1.3. Finding child participants

Finding participants for the studies in this research involved making contact with 3 new mainstream schools that had not previously worked with the University of Hertfordshire's Adaptive System Research Group, and re-contacting a school with specialist provisions for children with special needs. The first 3 studies were conducted by the principle investigator, with 50 children taking part in 100 interviews. The final study was conducted by an educational psychologist with the robot in a local mainstream school. Establishing and maintaining good working relationships with the staff at the schools was vital to ensuring the success of the research. All samples in this research were samples of convenience.

3.1.1.4. Recruiting assistance

One of the considerations when conducting the studies was to ensure that the children had equal minimal contact with the interviewers. Therefore, it was necessary to enlist the help of an assistant for each study, to take the children to and from the interviews.

3.1.1.5. Experimental setup

To ensure that the children were unaware that the robot was being controlled by a human operator, a control room was set up in each study where the robot could be controlled from via a remote connection. The control rooms consisted of a laptop remotely connected to the robot and a monitor with headset that received a live feed from one of the cameras recording the interviews. The control room setup for each study is described later in this thesis.

3.1.2. Ethical considerations

3.1.2.1. Ethical approval and prerequisites

Prior to conducting the research, each study was approved by the University of Hertfordshire Ethics Committee. One of the requirements for ethical approval to be granted was to obtain an Enhanced CRB Disclosure, therefore an application was made and an Enhanced Criminal Record Certificate was issued by the Criminal Record Bureau (CRB). In addition to the principal investigator, it was also necessary for the experimental assistant to have an Enhanced CRB Disclosure, therefore it was ensured that copies of the assistants CRB's were obtained and retained for the records.

3.1.2.2. Keeping the schools informed

The staff at the schools were kept fully informed about the purpose and structure of the study, and were also given copies of the CRB Disclosure for their records.

3.1.2.3. Obtaining participant/parental consent

Participant/parental consent was obtained before any studies took place. The consent form advised the participants/parents the purpose of the study and informed them of the particular arrangements for each study. The consent forms for some of the studies also gave an option of whether stills of the video footage could be used in scientific publications or not.

3.1.2.4. Well-being of participants

The studies were designed to be as stress free as possible for the children. If the children did not want to take part in the study they could cease the session at any point. In addition the children were given the opportunity at the end of each study to control the robot and see how the robot worked. The children were also given the opportunity to play with a large selection of other robotic toys during the post-study debriefs.

3.1.2.5. Participant privacy

The confidentiality and anonymity of the participants was kept at all times. No names or identifiable pictures were used at any time unless prior consent had been obtained.

3.1.2.6. Data security

All video data captured was stored in a secure location on 128 bit encrypted password protected hard drive. Each participant was assigned a unique ID and all corresponding data used this ID rather than the participant name.

3.1.3. Data analysis

3.1.3.1. Data sources

All of the interviews in the studies were recorded on three or four hi-definition video cameras to capture the verbal dialogue and physical behaviour during the interviews. A total of approximately 53 hours of video footage was captured from the first 3 studies (study#1: 21 hours, study #2: 26 hours, study #3: 6 hours). From this footage approximately 9 hours and 5 minutes were used for analysis (study#1: 04:18, study #2: 03:58, study #3: 00:49). In addition to video footage, questionnaires were also used to gather information about the children's perspective on various aspects of the interviews.

3.1.3.2. Measurements taken

The video data captured in the studies was processed by coding in the Observer XT software [68], and transcription. Various temporal aspects of the interviews were analysed, as were quantitative and qualitative aspects of the transcriptions. Additional methods of measurement also included questionnaire data, and independent blind coded engagement data. In total approximately 528 hours were spent on the measurement aspects of this research which was divided as follows:

Table 3.1.3.2.1: The amount of video footage gathered from each of the first three studies varied, as did the processing time for each aspect of the data gathered. This table gives an approximation of the processing time for each aspect of data collection in the three studies.

NOTE: TPM = Time Per Minute coding, NT = Neurotypical, SN = Special Needs

HH:MM	TPM		Study #1	Study #2	Study #3	Totals
Total footage	**NT**	**SN**	4:18	3:58	0:49	9:05
Total coding	30	20	129:00	119:00	16:20	264:20
Eye Gaze	17		73:06	67:26	N/A	140:32
Vocalisation	10	15	43:00	39:40	12:15	94:55
Basic gestures	3	5	12:54	11:54	4:05	28:53

- Approximately 264 hours spent coding 9 hours of footage in the Observer XT
- Approximately 100 hours spent transcribing dialogue
- Approximately 50 hours spent analysing transcriptions
- Approximately 2 hours spent processing questionnaires
- Approximately 12 hours spent by independent blind coders coding engagement data
- Approximately 100 hours spent on statistical analysis of the data, including developing an analysis system in excel

3.1.3.3. Reliability of data

To ensure the integrity of the results for this research, all data coded in the Observer XT was subject to an inter-rater reliability check. 20% of all of the videos coded were also coded by an independent coder following the coding scheme. The principal investigators coded videos were checked against the second coder's data to verify the reliability.

3.1.4. Expert feedback and actions

3.1.4.1. Gathering expert opinions

To establish the views and opinions of experts working with children that would potentially have a use for a RMI approach, contacts were formed with a UK Police Constabulary, an intermediary service and a range of other professionals who had relevant experience of working with children. Three separate user panels were arranged with the experts to give an overview of the studies conducted and the subsequent results, before asking the experts their views on this approach and what features they would require for such a system to be useful in their field.

3.1.4.2. Programming the robot

Once a set of requirements had been outlined by the panels of experts, the interface and control mechanisms were adapted to suit the outlined requirements. Implementing these changes before the field trial was necessary to make the system sufficiently useable by the educational psychologist testing the system in the final study.

3.1.5. Field trial with an expert using the system

In the final study the robot with a modified interface was used by an educational psychologist in a field trial. The purpose of the final field trial was to gather feedback and ultimately establish if a RMI approach could work for real world users. The key to conducting this study successfully was to arrange and co-ordinate times where the educational psychologist could test the system at the school where the study would take place.

3.2. Experimental setup

3.2.1. Procedure

The structuring of the first three studies involved a group pre-interview introduction, followed by individual interviews, concluded by a group debriefing. The group introduction was used to briefly familiarise the children with both KASPAR and the principle investigator at the school before the interviews commenced. During the introductory sessions we provided information on the purpose of the study (conducting interviews), and explained the nature of KASPAR (introduced KASPAR as a robot), however the term “interview” was not explicitly used during this introduction. It was emphasised to the children that they were not being assessed or graded on what they did or said in the interviews, and that there were no right or wrong answers. This was explained because we did not want the children to be worried or distressed about having the interview. Although the term “interview” was not used when describing the activity to the children, this work differs from the research conducted by Bethel et al [13], because the purpose of the activity was clearly focused on information acquisition without any other interactive tasks taking place. It was ensured that the children had equal minimal contact with both KASPAR and the human interviewer prior to the interviews commencing, as having disproportionate contact could adversely affect the results. During the introduction, it was explained to the children what we would be talking to them about, however the children were not provided with any precise details as to what they would be asked, as this could lead and influence what they might have said in the interviews. After all of the interviews had taken place, the children were given a group debrief to explain how KASPAR worked and given the opportunity to control the robot.

3.2.2. The robot

The robot used to conduct this research was a small child-sized, humanoid robot called KASPAR developed at the University of Hertfordshire. KASPAR has a minimally expressive face and arms that are capable of gesturing. This robot has been shown to be very effective when working alongside neurotypical children [66, 69] and children with autism [3, 4, 70, 71]. Robins et al. explored how children adapt to interacting with a robot and will mirror a robot’s temporal behaviour [62]. KASPAR has also been used to explore the role of gestures during Human-Robot Interaction [61], and how different types of embodiment affect interaction [66]. Previous research conducted with KASPAR would suggest that KASPAR is a suitable platform for this particular area of research investigating how children respond to a humanoid robot in an interview setting.

Figure 3.2.2.1: Images of the KASPAR robot used in this research

3.2.3. Hardware

KASPAR is a small stationary humanoid robot that contains 16 servos, which equates to 8 Degrees of Freedom in the head and neck, along with 4 DOF in each arm [2], see Table 3.2.3.1 for further details. The servos are controlled via a PC Interface to Dynamixel Bus. KASPAR is powered by a 12 volt lead acid battery that is capable of powering the robot for approximately 5 hours. The face of the robot is a flesh coloured silicon rubber mask from a child resuscitation dummy, which is produced by Laerdal [72]. The flexibility of the mask allows the servos to manipulate the mask to produce different facial expressions. KASPAR was initially designed to be a budget robot suitable for Human-Robot Interaction studies, therefore the majority of components for KASPAR are readily available off the shelf parts with the exception of the custom built metal parts that form the structure of the head, neck and joints. When using KASPAR, the robot is connected to a laptop with custom written software to control the robot via a USB cable. In an effort to make KASPAR seem as autonomous as possible in this research, all of the cables and hardware were concealed within a small padded box that KASPAR was placed on during the interviews.

Table 3.2.3.1: This table itemises the KASPAR robot's Degrees Of Freedom (DOF)

Head/Neck	Left Arm/Hand	Right Arm/Hand
Eyes Left/Right Eyes Up/Down Eyelids Open/Close Mouth Smile/Frown Mouth Open/Close Head tilt Left/Right Head roll Up/Down Head pan Left/Right	Shoulder rotate Up/Down Shoulder In/Out Arm rotate Left/Right Elbow Bend/Straighten	Shoulder rotate Up/Down Shoulder In/Out Arm rotate Left/Right Elbow Bend/Straighten

3.2.4. Software

KASPAR is controlled via a Java based GUI which can be customised for specific applications. To create behaviour/questions for the studies, a selection of poses integrated into the sequences were created in the software. First the sliders in the software are used to move the joints of the robot, see Figure 3.2.4.1. After the relevant poses have been constructed, the poses are put into sequences to create the behaviours (movements of the robot) to accompany the audio for the questions, see Figure 3.2.4.2. Once all of the behaviours have been setup, they are assigned keys, the behaviours (questions in the case of this research), can be activated by a key press, see Figure 3.2.4.3. During the first three studies, pre-programmed behaviours (including utterances of speech) were activated by the experimenter pressing designated keys. This approach is following the Wizard-of-Oz methodology (WoZ), widely used in Human-Computer Interaction (see Gould et al. 1983 [73]; Dahlback et al. 1993 [74]) and more recently has been used in HRI [75, 76]. For each study the program controlling KASPAR was specifically tailored with pre-programmed audio clips accompanied by appropriate sequences of movements. For each study two sheets with the speech phrases and corresponding keys were placed in the control room to aid the investigator in finding the correct key. The audio clips for KASPAR's voice were generated from text-to-speech synthesis software. Text-to-speech software was used rather than recordings of a natural human voice to maintain the robotic theme. A natural human voice coming from robot would most likely have impaired the perceived consistency of the robot in terms of appearance and behaviour, which has been shown to be important in the Human-Robot Interaction literature [77, 78]. In addition to this, using a synthesised voice helped maintain the distinction between the robot and the human interviewer. During all of the studies conducted as part of this research, the children were unaware that the KASPAR robot was being controlled by a human investigator prior to the interview taking place.

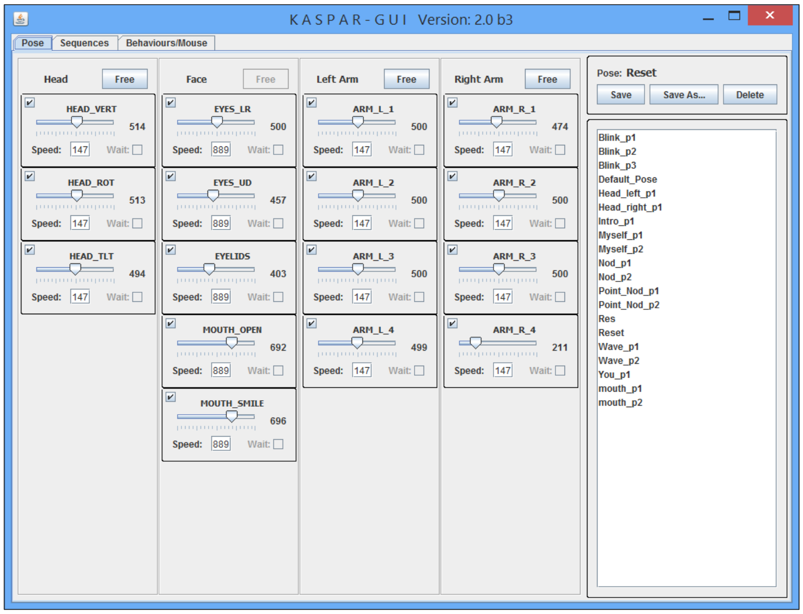


Figure 3.2.4.1: This is pose creation tab within the KASPAR software where the individual poses for the robot are constructed

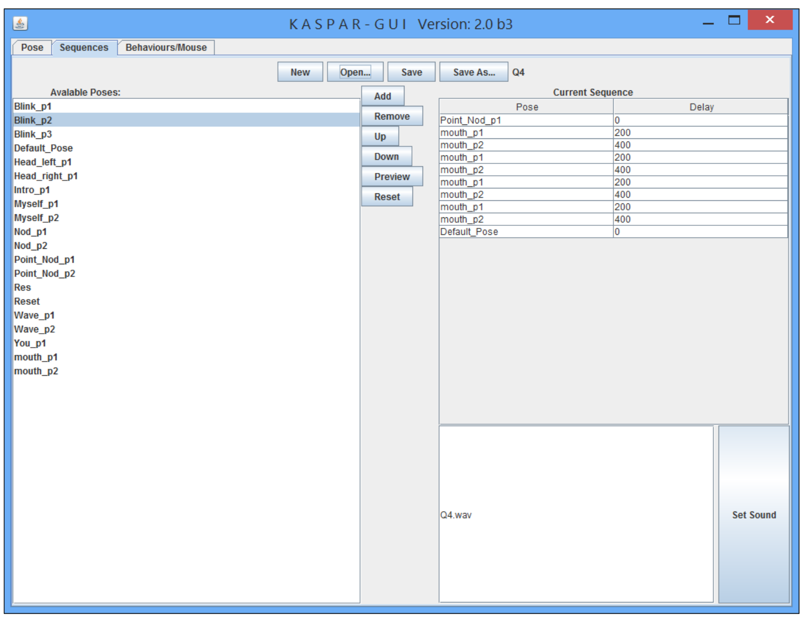


Figure 3.2.4.2: Sequence creation tab within the KASPAR where the sequences of movement are constructed from the poses

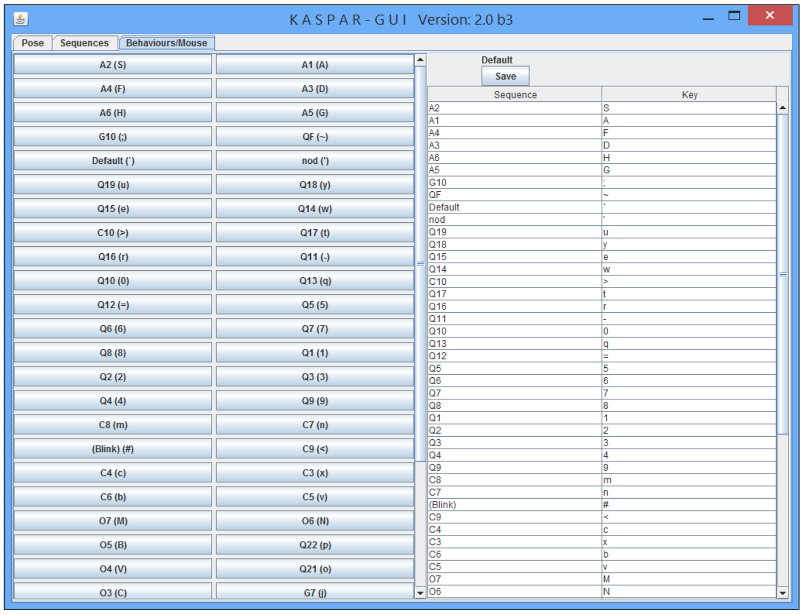


Figure 3.2.4.3: Behaviours tab within the KASPAR where the sequences of movement are activated

3.2.5. Body language

The ABE and research suggest that body language can play an important role during an interview [7],p.196. Based on the advice we received from the two specialist police officers from the Metropolitan Police and guidance from the ABE, the body language of the interviewer was kept minimal in both conditions. Because we wanted the interviews to be conducted in a similar manner to which a police interview would be conducted, it seemed logical to follow the advice of the ABE and specialist police officers.

3.2.6. Deciding how to measure

Before conducting any of the studies, it was important to carefully consider what and how to measure in the interviews [79, 80]. In an effort to find an effective means of measuring the interactions, inspiration was drawn from other HRI studies and research to establish what approaches have been shown to be successful previously. The evaluation of industrial and manufacturing robots is usually based on the speed, accuracy and reliability of the robot. The evaluation of social robots is much less clear and is often based on people’s subjective opinion. Kanda et al. [81] conducted an investigation where the body movements of participants were compared to subjective evaluations scores when interacting with a Robovie robot [56]. Kanda et al. explains that questionnaire-based methods are often “subjective, static, and obtrusive”.

Analysing human behaviours such as proximity, attitude, eye gaze and synchronised behaviours is more difficult to apply but produces more objective results. To gain a more complete picture Kanda et al. combined human behaviours with subjective questionnaires. It was found that people's subjective scores of the robot were positively influenced by well-coordinated interactions with the robot. This work illustrates how combining behavioural analysis based methods with traditional subjective questionnaire based methods can provide a more complete picture when evaluating how successful a robot is in its interactions, and what factors are pivotal in providing this success.

In addition to the work by Kanda et al. Tanaka et al. [17], used several means of analysis to investigate how engaged the children were and how their interactions with the robot compared to the interactions with their peers. The primary measure used in Tanaka et al.'s study was the coding of haptic behaviour, similar to psychology behavioural analysis based methods. Further to this method a continuous audience response method was used to subjectively measure the interaction quality. This novel method of analysis revealed that there was a strong relationship between haptic behaviour and the quality of the interactions. It also revealed that the high predictability of the second phase of the study negatively affected the quality of interactions between the child and the robot, but the interaction quality was quickly re-established once the full repertoire of behaviours had been restored. For measuring less straight forward behavioural responses this method could be adapted and used for the studies conducted as part of this research.

In light of the background research and practicality, the first three studies in this research would use multiple measures primarily derived from video footage that would later be coded in the Observer XT [68] and transcribed for a detailed analysis. Using video coding and transcriptions would allow for a large amount of information to be collected about different aspects of the interviews, helping to create a complete quantitative and qualitative picture of the interactions being analysed. In addition to the coded and transcribed data, supplementary measures from questionnaire data and demographic data were also used to assist the analysis. Bethel et al. emphasises the importance of using multiple means of evaluation when HRI conducting research [80], which is a recommendation that has been followed whilst conducting the research in this thesis.

3.2.7. Counterbalancing participants

In order to compare the two conditions, each child experienced two interviews, one with KASPAR and one with a human experimenter. The interviews followed the same structure on both occasions. A two-phase counterbalancing method was implemented to reduce the chance of the interview order adversely influencing the results. Half of the children were interviewed by KASPAR first and half were interviewed by the principle investigator first. Counterbalancing was also applied in terms of gender and year group, so that each group had the same number of boys/girls and year groups. Assignment to each of the two groups was otherwise random for any particular child.

3.2.8. Equipment and resources

Prior to conducting the experiments for this research, it was important to consider the quality and layout of the equipment, as these factors can have an impact on the accuracy of data and results achieved. When setting up equipment for the studies, careful consideration was given to what would be measured to ensure all the necessary information was captured. Because the main measures in the first three studies would rely on the coding of audible and visual information, as well as transcriptions, it was important to ensure that the cameras were positioned effectively. For example, when measuring eye gaze it is important that the camera is focused on the primary subject from an appropriate angle with a clear unobstructed front view of the primary subjects face and also has the object(s) or subject(s) in view to which the eye gaze is being coded, in the case of an interview, the interviewer. It is also important to consider how much the primary subject is likely to move. For the three studies conducted as part of this research, we decided to use 3 to 4 cameras, capturing video data from several different angles. This would allow us the option to view the video from a different perspective and would also assist if any obstructions occurred to the view of the primary coding camera. When planning the studies it was decided that 2 cameras would be focused on the child’s face from behind the interviewers on opposite sides, whilst 1 or 2 cameras were used to record the interviewers face and the overall activity in the room. In addition to setting appropriate camera angles, the cameras were set to boost the microphone pickup as the audible information would later be transcribed from the recordings.

3.3. Measures and methods employed to acquire data

3.3.1. Video coding

3.3.1.1. Design of coding schemes

When designing the experiments it was vital that the primary points of measurement were outlined before the studies were conducted, as this is the basis for defining what will be coded from the video, and consequently would affect where the cameras should be positioned in order to capture the data. To ensure all the immediate relevant information was collected, a baseline coding scheme was drafted which could have elements added to or subtracted from, dependant on what was useful for the given study. To ensure clarity of the coding scheme, each element or behaviour being coded consisted of a definition, measurement type, purpose and coding time.

3.3.1.1.1. Definition

Setting a clear definition minimises room for misinterpretation and enables a second reliability coder to follow a strict rule. I.e. eye gaze towards the head of the experimenter.

3.3.1.1.2. Measurement Type

Deciding what measurement type to use depends on the behaviour that is being analysed and the data that is being collected. For example, if examining how interested a child is in a particular aspect of a toy, it may be useful to code how many times the child pushes a particular button on the toy. This would be coded as a point event to count the number of occurrences. In a RMI setting, we considered that it would be useful to investigate how long the child looks at the interviewer, as this may be an indicator as to how interested the child is in the interviewer. Therefore, coding eye gaze towards the interviewers head would be one effective means of capturing this information and would be coded as a state event to capture the duration.

3.3.1.1.3. Purpose

When defining a coding scheme it was useful to explain the purpose of the behaviour that was being coded. Stating the purpose of coding the behaviour is particularly useful when explaining to other researchers why the behaviour is important to code. In this research, outlining the purpose of coding the behaviours served as a useful means of explaining to the project supervisors exactly why the behaviour was being coded. In other projects the purpose can also be used to assess the importance of coding a particular behaviour, and if there are time or budget constraints, the coding scheme can be reduced to focus on the most important aspects of the interaction.

3.3.1.1.4. Coding time

Defining strict rules for when behaviours should be coded is one of the most important aspects of the coding scheme for both making an accurate measurement and acquiring strong inter-rate reliability. The coding schemes used in this research clearly defined precisely when the coding for each behaviour should begin and end for a state event, and at what particular moment for a point event. For state behaviours, regardless of the behaviour being coded, the times at which coding begins and ends were similar. When creating coding schemes for the research in this thesis, this rule was strictly applied.

3.3.1.2. The standard coding scheme

Before conducting the studies, a standard video coding scheme to be used in the Observer software was developed to measure specific aspects of the interactions, primarily focusing on the vocal aspects of the interaction. The base coding scheme that was used for the first three studies consisted of the behaviours outlined in Table 3.3.1.2.1.

Table 3.3.1.2.1: Default coding scheme used to define behaviours to be coded

Eye gaze duration (Mutually Exclusive): KEYS – Start/Stop = E
DEFINITION – Eye gaze is when the child is looking towards the interviewers face
TYPE – State Event to acquire the duration of the eye gaze
CODING TIME – Coded from first contact with interviewers eyes/face contact to last contact before looking away

Vocal & Response durations (Mutually Exclusive):
Start/Stop gap KEYS – Start/Stop = #
DEFINITION – Time at beginning and end of interviews
TYPE – State Event to assist mutually exclusive coding
CODING TIME – Coded to the first word of the interview and from the last word of the interview

Child speaking duration KEYS – Start/Stop = C
DEFINITION – Time child spends responding to a question or response from the interviewer
TYPE – State Event to acquire the duration of the child's answers
CODING TIME – Coded from when the child speaks to when they finish speaking. Any gaps greater than 2 seconds will be coded as Child pause

Interviewer speaking duration KEYS – Start/Stop = I
DEFINITION – Time interviewer spends asking a question or responding to the child
TYPE – State Event to acquire the duration of the interviewers questions
CODING TIME – Coded from interviewers first word to last word

Child pause duration KEYS – Start/Stop = P
DEFINITION – If a child pauses for 2 seconds or more in-between their response, this will be recorded as a pause i.e. Child > Child. This will be used for both logging the amount of pause time but also amount of eye gaze in theses pauses
TYPE – State Event to acquire the duration of the child's pauses
CODING TIME – Coded from child's first pauses the beginning of the next word when they resume speaking

Response time (Interviewer>Child) KEYS – Start/Stop = i
DEFINITION – Time between the end of the interviewer's verbalisation and the beginning of the child's verbalisation
TYPE – State Event to acquire the duration of the gap
CODING TIME – Coded from the last word of the interviewer to first word of the child

Response time (Child>Interviewer) KEYS – Start/Stop = c
DEFINITION – Time between the end of the child's verbalisation and the beginning of the interviewer's verbalisation
TYPE – State Event to acquire the duration of the gap
CODING TIME – Coded from the last word of the child to first word of the interviewer

3.3.1.3. Output of coding measures

The outputs for the standard coding measures were as follows:

Interview duration – The full duration of the interview from start to finish. This was used to assess if there was any difference in the time the interviews would take.

Eye gaze duration – Defined as the child looking towards the interviewer's face. Measuring eye gaze duration would allow us to evaluate the different amounts of eye gaze towards the robot compared to the human. Eye gaze would be measured proportionate to the duration of the interviews, this is because the interviews would inevitably vary in length and it is important to take this into account.

Child response duration – Total amount of time the child spends speaking to the interviewer throughout the full duration of the interview.

Child pause duration – Total amount of time the child pauses for 2 seconds or more whilst speaking throughout the full duration of the interview.

Interviewer response duration – Total amount of time the interviewer spends speaking to the child throughout the full duration of the interview.

Response time child > interviewer – Total amount of time throughout the full duration of the interview that the interviewer takes to respond to the child.

Response time interviewer > child – Total amount of time throughout the full duration of the interview that the child takes to respond to the interviewer.

Because of the exploratory nature of this work, additional details of the interactions were originally incorporated into the coding scheme, behaviours such as nodding and shaking of the head, shrugging of shoulders and smiling. Literature suggests that body language can provide useful indicators about people's thoughts and feelings [82], however these behaviours were not used in the final dataset because of their infrequent and inconsistent occurrence. In addition to this, behaviours such as smiling can be very difficult to code due to their being scope for misinterpretation. One of the most accurate ways to code this type of information would be to adopt the FACS system (Facial Action Coding System) [83]. Measuring the aspects of the interactions that we selected would serve to generate sufficiently rich data on the interactions during the interviews. Measuring the response times between the interviewer and the child was useful and essentially free data. Because the times of the interviewer and the child needed to be coded, the gaps between easily revealed themselves in the coding without much additional effort, therefore when coding vocalisation between two people it makes sense to automatically include this into the coding scheme.

3.3.1.4. Coding best practice and rigour

When coding the videos for this research several best practice rules were applied in order to maintain a high level of accuracy and consistency. The rules applied and the reasons for their implementation are discussed in the following subsections.

3.3.1.4.1. Eye gaze

Eye gaze is a difficult behaviour to code well, therefore when coding eye gaze, strict rules were applied and followed. When coding eye gaze, we are often looking for gaze directed toward a particular area, whether this be an area of person or object. Eye gaze should always be coded as a state event, because some individuals will direct their eye gaze for long periods, while others will look for lots of very short periods. The only way to confidently establish the amount of eye gaze directed toward a particular area is to measure the entire duration. Eye gaze should only be coded whilst the eye gaze is focused on the area of interest. Therefore coding would take place from the first frame that the eye gaze is focused on the area to the last frame before looking away from the object, in the case of this research the interviewers face. When coding eye gaze it is important to run the video at a slower rate, usually half of the normal playing speed. This is essential as accurately coding eye gaze at full speed is impossible. Coding eye gaze is directly affected by the camera angle that the experiments are recorded from. Eye gaze can only be coded when the eyes of the participant and the area of interest are in view. If the eyes of the participant are obstructed from view, or the area of interest are not in view then this cannot be coded, this is why it is essential that camera angles are carefully considered before carrying out an experiment. Measuring eye gaze is a difficult and time consuming task, but can provide useful information in HRI studies, particularly when trying to establish how interested and engaged an individual is in a robot.

3.3.1.4.2. Audible information

The way that audible information is processed depends on the type of audible data. For example, if working with individuals that make noise it may only be necessary to code how long these vocalisations are made. However if the participants are speaking full words and sentences, then it may be more useful to transcribe what is said and analyse these transcriptions. For the research conducted in this thesis it was useful to both code and transcribe the audible information, to establish if there was any difference in how much a child said and how long they took to say it. Measuring both would assist in analysing the difference between a Human-Robot Interaction and a Human-Human Interaction.

When coding audible information it can be a lot more difficult to be specific and accurate, particularly if there is background noise or if there is an overlap of speech with more than one person speaking at a time. Audible information was coded at half the videos original speed as this was the most accurate way of coding. In the studies conducted, the majority of the time there was only one person speaking at a time. However there was the occasional overlap of speech which could make it difficult to differentiate between voices when being played in slow motion, therefore it was important to pay close attention in these cases.

In addition to the basic coding of when an individual was speaking, it was also useful to code the response times, particularly as the interviews were a dialog interaction. Examining this detail would allow us to analyse the difference in the communication between the child and the robot, and the child and human. Investigating response times in interactions can be useful for understanding interaction dynamics.

3.3.1.4.3. Inter-rater reliability

When conducting research it is important to verify that the results are real and replicable by other researchers [84]. In this research all video coded data was verified by a second independent reliability coder. 20% of the videos were selected for second coding, and the selection of this material was carefully considered. When selecting videos for the second coder to code, a counterbalance selection was chosen to represent a wide range of videos from the dataset. Criteria such as experimental condition, gender and age group were used when counterbalancing.

Before coding data from experiments that will be used for the basis of scientific evidence, it is important to give new coders appropriate training and testing. To ensure coders have an understanding of good coding practice, all video coders that worked on this research were required to undertake training on test data. The training video that the second coders were required to code before commencing work on real videos contained audible and visual aspects that were to be coded. For example, coders needed to code eye gaze and speaking, similar to the coding scheme used in this research in order to practice and test their coding ability prior to commencing coding on real data. The results of their coding were checked for reliability against an experienced coders coding of the video and provided that the reliability was over 0.80, the coder could commence work on the main video. This testing ensured that all coders met a particular standard, and the data that they produced was likely to be reliable.

In this research the coding scheme was clearly defined, and the importance of defining a clear and concise coding scheme becomes apparent when a second independent reliability coder begins to code videos. Clear definitions and coding times are a crucial factor for achieving reliable results. It was equally important to stick to the rules of the coding scheme once the coding begins. The more the rules of the coding scheme are bent or broken, the more likely the data produced is to be unreliable, and in turn the kappa score achieved by the inter-rater reliability analysis is likely to be lower, indicating that the data is not as reliable. It is therefore very important that the coding rules are adhered to as closely as possible. In this research the inter-rater reliability scores were generally very high, indicating that the coding schemes were sufficiently detailed and followed explicitly.

3.3.1.4.4. The impact of resources on results

When coding data, the resources that are being used can affect the accuracy and the speed of the coding. Factors such as screen size, PC power, video quality and file formats all play important roles when coding video data. For this research it was ensured that the video data was of sufficient quality to accurately code the behaviours being analysed. When analysing behaviours such as eye gaze, it is particularly important that the quality of the video is sufficient to accurately establish where the eye gaze is being directed. It was also equally important to have a sufficiently sized screen to view the videos when coding. Using a large screen will not only allow a coder to make a better informed more accurate judgement, it can also increase the speed at which data can be coded, as it is much easier to see behaviours on a larger screen. Other factors such a PC power and file format can also significantly affect the rate at which data is coded. Having a slow PC that lags and takes a long time to respond will adversely affect the speed at which data can be coded. The file format of the video can also have implications on the speed at which a video can be coded. To ensure these potential issues were not encountered, a check was performed to see how the videos played at full, half and a quarter speed in both the normal playing direction and reverse before coding. Some file formats will not play well in reverse or at slower rates because of the way they are encoded. It was therefore important to check this before beginning coding.

3.3.2. Transcription analysis

Due to the nature of this research, a good portion of the transcription analysis would involve both quantitative and qualitative investigation. To ensure consistency throughout the duration of this research a standard analysis system was created in excel that would be applied to all three studied. To perform a detailed analysis of the transcriptions of interviews in the studies, a series of excel spread sheets were setup and formulated to process the information both in terms of content and length. The responses in the interviews were broken down by response. For example, the interviewer would ask a question then the child would respond, each response had its own cell which corresponded to the question ID. Breaking questions down by ID would allow for detailed analysis for each question. When transcribing videos it is easy to make mistakes, therefore all videos were transcribed then the transcriptions were checked back against the videos, and any mistakes were rectified before analysis. The general means of analysis applied to the transcriptions were as follows:

Word count – Total number of words spoken by the child throughout the full duration of the interview excluding filler words. This would be used to measure how much the children spoke in each interview.

Filler word count – Total number of filler words spoken by the child throughout the full duration of the interview. People often use filler words such as “err”, “errm”, “hum”, etc. and these words were included in the raw transcriptions. When analysing the transcriptions for a word count these filler words were not counted in that analysis but a separate filler word analysis would be performed.

Proportionate word count – Total number of words spoken by the child throughout the full duration of the interview excluding filler words proportionate to the total number of words spoken by the interviewer throughout the full duration of the interview.

Key-word count - Total number of key-words spoken by the child throughout the full duration of the interview. The key-words chosen were different for each study and related to the questions that the children would be asked.

Key-points – During each of the studies we also investigated the transcriptions from a qualitative perspective by performing a key-point analysis. The key-points would be different for each study and similar to the key-words, would relate to the questions. However a key-point would be a specific piece of information which was acquired by reading the transcriptions and logging the key-points. This information would be analysed both qualitatively and quantitatively. The qualitative aspects would be the specific details of what the children were saying and the consistency of the information between the two interviews. The quantitative aspect is the numerical logging of each specific piece of information given by the child and the statistical analysis of the logged information. The latter would be done in order to understand how many key-points the child revealed.

Each of the means of analysis for the transcription data could also be broken down by question to investigate the specifics of each question. Having this ability would be useful for analysing how the children would respond to different types of question from each interviewer.

3.3.3. Questionnaire data

In addition to the coded data and transcriptions, a simple standard questionnaire was also developed for the children to complete immediately after each interview. The purpose of this was to establish what the children thought of the whole experience and in particular what they thought of KASPAR. In order to ensure the children were not overwhelmed, the questionnaire was kept short and simple, see Figure 3.3.3.1. The information that the questionnaire was designed to reveal was:

- How interesting they found the experience
- How difficult they found the interview
- How much fun they had participating
- How long they thought the interview took

Talking to KASPAR/Luke was?				
1 Very Boring	2	3	4	5 Very Interesting
1 Very Hard	2	3	4	5 Very Easy
1 No Fun	2	3	4	5 Fun
1 Taking a long time	2	3	4	5 Very quick

Figure 3.3.3.1: Standard post-interview questionnaire used in studies 1 to 3

3.3.4. Demographic data

In addition to the data relating to the interviews, demographic data was also collected about the children participating in the study. The data collected included: age, gender and year group. In the third study additional details of the children's special needs were also gathered from the teacher. This data would give us the option to analyse if these factors had any impact on how the children interacted with the different interviewers during the interviews.

3.4. Analysing and tests performed on data

The initial results gathered from the data were basic information such as: ranges of data, mean and median values, standard deviations and confidence intervals. This was followed by testing the data for normality with a nonparametric one-sample Kolmogorov-Smirnov test. This test was performed on the difference between the robot and the human interviewer conditions because normally distributed data would allow us to use the confidence intervals as a measure of the true differences that could be expected between a robot and a human interviewer. If the data is not normally distributed however, an assessment using median values and ranges complemented by boxplots would be more appropriate. Testing the data for normality was an essential step as this would indicate which further tests would be appropriate to perform on the data.

Chapter 4: How do young children respond to a humanoid robot in an interview setting (Study #1)

4.1. Introduction

Robots in therapy and education are typically meant to facilitate learning and/or therapeutic changes in the children. In this proposed novel application area, robots would be used as mediators between a professional human interviewer and a child, providing a simple and enjoyable interaction partner with the purpose of engaging the children in the interview for the retrieval of vital information. Exploring the possibility of using robots to interview children may reveal whether RMIs could be a valid addition to existing methods of interviewing children by professional staff such as police officers, social services staff or educational psychologists. However, before commencing investigations in the sensitive areas of interviews with children in a social services or police context, we need to establish whether or not a humanoid robot is, in more general terms, acceptable as a robotic interviewer, e.g. will children take the interviews "seriously", i.e. discuss factual information, as opposed to treating the situation as an entertainment activity where they use their imagination? This study established baseline using a quantitative experimental approach.

This chapter outlines the first experimental study in this thesis examining how children's responses towards the humanoid robot KASPAR in an interview context differ in comparison to their interaction with a human in the same setting. Twenty-two children aged between 7 and 9 took part in this study. Each child participated in two interviews, one with an adult and one with the humanoid robot KASPAR. Measures include the behavioural coding of the children's behaviour during the interviews, transcriptions from the interviews and questionnaire data. The questions in these interviews focused on a special event that had recently taken place in the school. The results reveal that the children interacted with KASPAR very in a similar manner to how they interacted with a human interviewer. The quantitative temporal analysis reveal that the most notable difference between the interviews with KASPAR and the human were the duration of the interviews, the eye gaze directed towards the different interviewers, and the response time of the interviewers. The results from this study illustrated that using a robot such as KASPAR in an interview setting seems to work and was worth further investigation.

4.1.1. Research questions

The purpose of this study was to investigate two research questions that we identified as the first necessary steps in establishing whether or not a robot can be used as an interviewer for young children which were: "RQ-1: To what degree do children respond to a robot in an interview setting?" and "RQ-2: How do children respond to a robot compared to a human interviewer?". Ascertaining if children will respond to questions from a robot in a formal interview setting was the purpose of this study. In order to answer the first two research question we choose to examine how the children's non-verbal and verbal behaviour would differ in the two experimental conditions. We also examined the children's opinions about the interaction and in particular, the interaction partner. In addition to this, because interviews are often about information recovery, we also analysed the content of the children's responses, to examine the information that they would disclose to each interview partner and see if there was any difference in the information that they would provide.

4.1.2. Hypotheses

In this study we expected that the children would be more interested in KASPAR as a novel object [85], and would direct more behaviours that indicate interest (e.g. eye gaze) towards the robot compared to the human interviewer. In terms of information recovery, on the one hand, one may expect that children would talk more and reveal more information to the human experimenter, since the children are very used to the situation of being asked questions by a human (e.g. at home or at school) rather than talking to a robot. On the other hand, if the children experience the robot as an enjoyable and comfortable interaction partner (compared to the human experimenter who is a stranger to them) then they might disclose more information to KASPAR, thus a clear preference for either the human or robotic interviewer was expected.

4.2. Methods

4.2.1. Participants

The study was conducted in a local UK mainstream primary school in Hertfordshire with children aged between 7 and 9 years old (UK year groups 3 and 4). The study involved 22 children, 20 of which produced useable data[1]. Of the 20 children 10 were in year 3, 10 were in year 4, 10 were female and 10 were male. The majority of the children in year 3 were female and the majority of year 4 were male. Of the 20 children, 3 have been diagnosed with 'some form of autism (according to the teachers)[2]. In accordance with the ethics procedure, informed consent was obtained in writing from the parents of all the children participating in the study as outlined in Section 3.1.2.3. The ethics protocol number for this study is 1011/161, the relevant documentation can be seen in Appendix 2.1. The adult interviewer was the author of this thesis.

4.2.2. Procedure

The interviews that we conducted with the children took place on four days over a two-week period. As outlined in Section 3.2.1, the children were briefly given a group introduction to both KASPAR and the human interviewer at the school one day before the interviews commenced. In this introduction, it was explained that we would be talking about the Red Nose Day[3] talent event that had recently taken place. However the children were not provided with any details of what they would be asked as this could lead and influence what they might have said in the interviews. After the group introduction had taken place the children were split into two groups with each child experiencing two counterbalanced interviews, one with KASPAR and one with a human experimenter. The counterbalancing of the children was performed in accordance to the specification outlined in Section 3.2.7. The interviews were held one week apart with the same interview structure being followed on both occasions. The children in group 1 were interviewed

[1] The data from the first interview conducted was not consistent with the other interviews in this study and was therefore considered a pre-study pilot. There was a technical difficulty with the robot in one session which meant the session had to end early therefore we could not include the data from the sessions with this child.

[2] The results from the sessions with the children with autism were consistent with the results of the neurotypical children, we therefore decided to include the data from these sessions in the dataset for this study.

[3] Red Nose Day is a bi-annual national event in the UK to raise money for charities, http://www.rednoseday.com/.

by KASPAR first, there were 6 males and 5 females[4], the average age in this group was 8. In group 2 the children were interviewed by the human first, there were 5 males and 6 females; the average age in this group was 8.3. The primary units of analysis were verbal communication, eye gaze and information disclosure as described in Section 3.3. After the interviews had taken place the children were given a group debrief to explain how KASPAR worked, and also given the opportunity to control the robot.

4.2.3. Experimental setup

The study was conducted in a small room with a recessed portion that was mostly hidden from the children. The interviews took place in the main large area of the room at a table, while the recessed part of the room was used for the robot control tent. We used a small tent to fully hide the controls and monitor of KASPAR as the partition alone would not have fully hidden the equipment and controller. The control tent housed a small monitor with a wireless connection to camera #1 for viewing and listening to the children, and a laptop that controlled KASPAR via a remote connection. This was essential, as we needed to know what the children were saying in order to make KASPAR respond appropriately. In accordance with the pre-study considerations outlined in Section 3.2.8, camera #1 was behind the interviewer to the left and camera #2 was also behind the interviewer and to the right, both of these cameras were recording the front of the children to capture eye gaze, while camera #3 was recording the front of the interviewer.

Both KASPAR and the human interviewed all the children on two separate occasions one week apart. The interviews were always led by the principle investigator in person or remotely via KASPAR. This was important to maintain consistency, making sure that the responses and questions from both KASPAR and the human were the same. The children were taken to and from the interviews by a second female researcher unknown to the children. This second researcher remained in the room during the interviews in case of any technical difficulties with the robot, but was as non-reactive as possible in order to avoid interfering with the experiment. Immediately after the interview, the children were asked to complete a questionnaire and post it into a box located on a separate table. The second researcher answered any question's that the children had about the questionnaire. The layout of the room and experimental setup is shown in Figure 4.2.3.1.

[4] One of these females was the child not included in the final dataset because of technical difficulties.

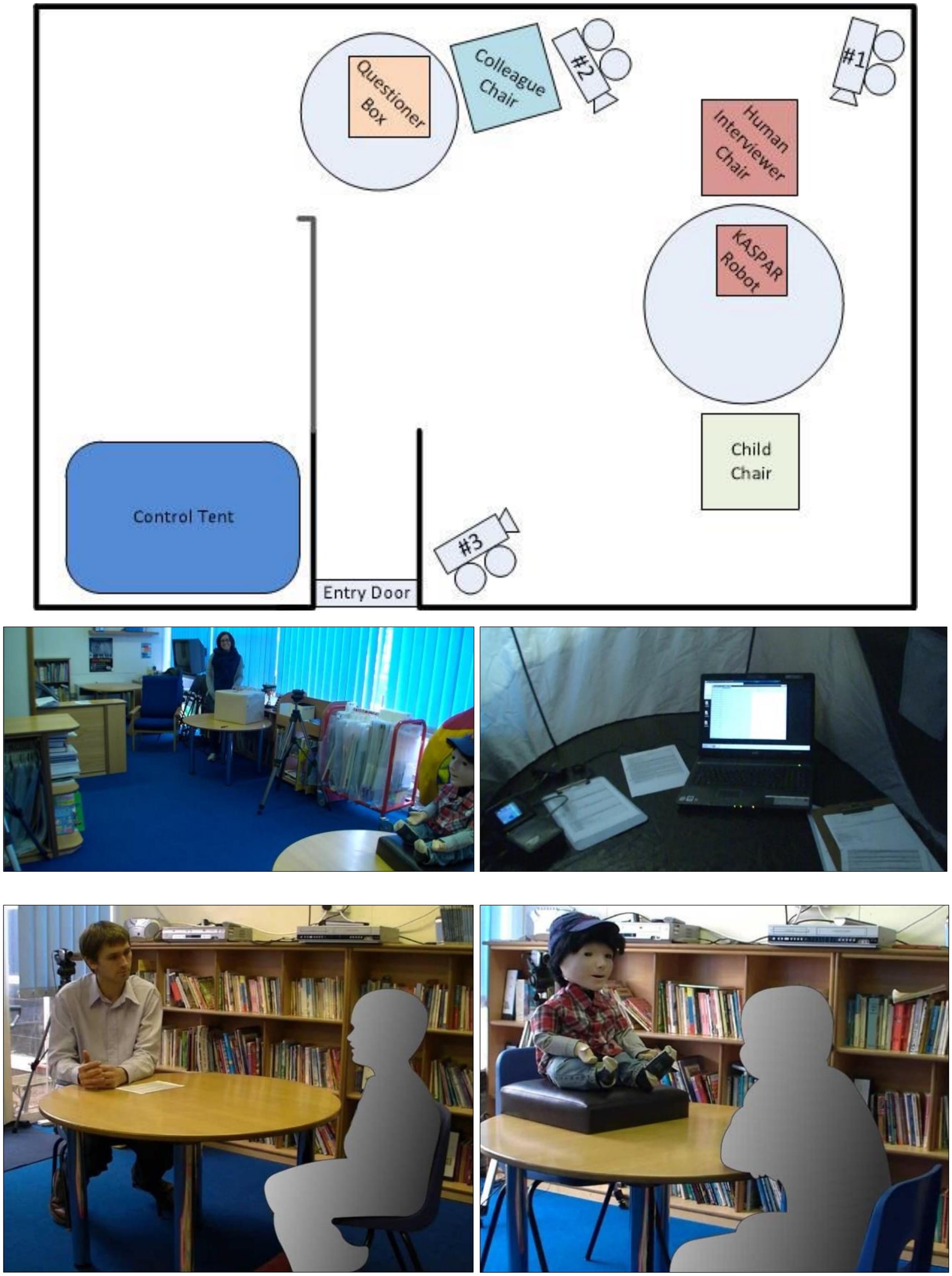


Figure 4.2.3.1: An overview of the room layout and interview setup

4.2.4. The interview

The interviews began with a short introduction of getting to know each other's name and ascertaining other general details such as the child's age, if they have any siblings etc. These questions were easy for the children to answer and helped to establish a rapport for the rest of the interview. We then proceeded towards the main topic, the talent event that the children had been involved in. Moving towards the main topic was achieved by asking the child "what are we going to talk about today". If the child did not remember they were reminded that they were there to talk about the talent event. Research and practice have shown that the most detailed and reliable answers are secured from open questions [18],p.27, therefore the majority of interview questions were open questions. This maximised the children's freedom to express themselves and might indicate who/what the child is more comfortable with based on how much they say and what they would say.

The questions that focused on the main topic varied in difficulty and this was reflected in the answers that the children gave. For example, the children found the question about who won the event much easier than the question about the judges who took part in the event. Almost all of the children correctly named both of the winners of the event. However, there was a much greater variation in the number of judges that the children could remember. This is possibly because during the event there would have been much more focus on the winners. Also, one of the judges of the event was unfamiliar to the children and it may have been harder for them to remember the name of this individual.

The questions in this interview primarily focused on facts, similar to how the police would conduct an interview. When the police or social services are trying to gather information from an individual they are interested in the facts of an event, as it is these facts that will be used to establish what has happened before deciding on what action to take [7],p.49, p.56, p.229. In addition to this, when making a prosecution, it is the facts and key-points that are used to make the case, rather than the feelings of the individual, because without the facts and points of proof a prosecution cannot be made.

The interviews concluded by thanking the child for their time and saying that it has been nice talking to them. In these interviews we adhered to a rigid structure with set sayings in order to compare the two different conditions. The majority of the structure and questions for the interview were derived from the guidance of the ABE document as stated in Section 2.1. A sample of the questions used can be seen below, a full list of interview questions is provided in the Appendix 2.2. Note: the main interview questions in Appendix 2.2 are listed on the order in which they were asked.

- Tell me about yourself.
- Tell me what we are going to talk about today.
- Describe the event to me.
- Tell me about the judges.
- Explain what happened in the final on Friday.
- Describe for me what the winner got.

4.2.5. Measurements

The questions and interview structure were reviewed and revised several times before trialling the structure, setup and equipment in the laboratory at the University with adult volunteers. In the schools data was collected from three cameras that recorded the interviews. In addition to the data from the recording of the interview, the children were also asked to complete a questionnaire immediately after each interview. This questionnaire was the standard questionnaire outlined in Section 3.3.3 to establish what they thought of the whole experience and in particular what they thought of KASPAR. Once all the interviews were complete, the video footage was transcribed and then coded using the Observer XT software [68]. We used the generic coding scheme outlined in Section 3.3.1 to measure eye gaze from the child to the interviewer, verbal communication both in terms of spoken words, duration of responses and gaps between responses from the child and the interviewer. Other body language such as nodding and shaking of the head was also measured but were too sparse and varied to provide useful information.

4.2.5.1. Communicative content

A full transcription analysis as stated in Section 3.3.2 was also used in this study to analyse the transcribed data. The key-words and key-points for analysis related to the questions asked in this study. The key-words chosen specifically focused on four areas:

- Family members (brothers, sisters, etc…)
- Names of judges for the talent event
- Prizes for the winners of the event
- Names of the event winners

The questions in this study were designed to recover information about 6 main categories. Each category had a specific information criteria defining it as a key-point. The key-points from the transcriptions consisted of:

- Family
- Pets
- Event acts
- Judges
- Winners
- Poster activity

Key-points - Family category

Since one of the questions in the interviews asked about the children's siblings, in this category we analysed how many family members the children mention in total throughout the duration of the interview, and how many family members they state by name. We also compared the names given in both experimental conditions, to establish the consistency of the facts disclosed in both interview conditions.

Key-points - Pets category

One of the introductory rapport building questions related to pets. Similar to the family category, we analysed how many pets the children mention and how many pets they stated by name for both experimental conditions.

Key-points - Event acts category

The questions in the interviews were designed to acquire information about the event the children either took part in or which they had witnessed. In this category we logged the number of types of acts that the children mentioned, the number of acts in the event, the number of people named that took part in the event, and a comparison of the names to check consistency. With regards to the types of acts this could refer to dancing or singing for example. If a child stated a year group and an act this was also counted as a type of act because this would be a specific type of act. The number of acts in the event refers to how many acts the child stated. For example the child may have said that there were 4 types of act (dancing, singing, acting, and magic tricks) but only referred to 2 acts that were performed (the winners and the chosen act from their own year group). In addition, the number of names the child mentioned who were in an act in the event was also logged.

Key-points - Judges category

Since the questions were based around a talent event with judges, we specifically asked questions to gather information about the judges. By analysing the transcriptions of what the children said we established how many judges there are in the event and also record how many of the judges they name, again with a comparison for consistency.

Key-points - Winner's category

Because the event was a competition there was one winning act with two children in the act. Some of the questions in the interview were designed to find out about the winner and what they received. We logged if the children could remember the winners name and if they were aware of the prize that the winners received.

Key-points - Posters category

Because there was only one act selected from each year to participate in the main talent event, many of the children were involved in a poster making activity to support their class mates. Because this seemed to be a significant and important part of the event for many of the children, we also recorded how many of the children mentioned this activity.

4.3. Results

4.3.1. Data analysis

In accordance with Section 3.4 we compiled the data and performed a series of nonparametric one-sample Kolmogorov-Smirnov tests on the data acquired for different measurements during the experiment. In addition to this we also created boxplots to assist in visually analysing the data.

4.3.1.1. Temporal analysis

The results of the analysis found that interviews with KASPAR lasted much longer, with a median difference of 1:56 (minutes : seconds), indicating that interviews with KASPAR took longer. Table 4.3.2.1 shows the average length of the interviews with KASPAR were approximately 6:26, whereas the interviews with the human were approximately 5:02. However, there was considerable variation in the durations of the interviews with both interviewers as Figure 4.3.3.1 illustrates. The interviews with KASPAR ranged from 3:44 to 10:45 whilst interviews with the human experimenter were between 3:24 and 7:33. The reason for the differences in the interview durations can be found by delving further into the data. Table 4.3.2.1 shows a breakdown of the response durations and response times for both the child and the interviewer. In particular we found that KASPAR took much longer to respond to the children than the human interviewer due to the technical limitations of the system. Figure 4.3.3.6 illustrates that on average throughout the full duration of the interview KASPAR took approximately 1 minute longer to respond to the children than the human interviewer who took approximately 20 seconds to respond to the children throughout the full duration of the interview. Further to this there was also a difference in the response times of the children to the interviewer. Throughout the full duration of the interviews the children took approximately 10 seconds longer to respond to KASPAR than the human interviewer, see Figure 4.3.3.7. With regards to the children's response duration, on average the children spent an additional 32 seconds speaking to KASPAR. However, the time which both of the interviewers spent speaking was very similar, see Table 4.3.2.1 and Figure 4.3.3.5.

Figure 4.3.3.2 shows in the majority of cases the children appear to look towards the face of KASPAR more, however given that the interviews with KASPAR lasted longer this is almost inevitable. To compensate for the length of the interviews, the results were normalised and calculated relative to the interview duration. Table 4.3.2.1 again indicates that the children generally look towards the face of KASPAR more that the human interviewer, however there are some outliers that either look at KASPAR a lot more or a lot less than the human interviewer proportionately, see Figure 4.3.3.3.

4.3.1.2. Textual content analysis

When analysing the amount of words that the children spoke to the robot compared to the human interviewer we found a median difference of 28.5 indicating that the children would use more words when speaking to KASPAR than the human interviewer, see Figure 4.3.3.9. Table 4.3.2.2 shows that the mean number of words that the children spoke to KASPAR was 363 words and 329.5 words to the human interviewer. In addition to this the amount of words the children used relative to the amount of words the interviewer used is very similar, as can be seen in Figure 4.3.3.10. The analysis of the filler word count found that the number of filler words the children used was very similar in both conditions, see Figure 4.3.3.11. On average there was just 1 words difference in the amount of filler words the children used between each condition. However, the number filler words used with KASPAR ranged from 2 words to 101 words, and with the human experimenter from 2 words to 63 words, see Table 4.3.2.2. The amount of key-words that the children used with KASPAR compared to how many they used with the human interviewer was also very similar, see Figure 4.3.3.12. Table 4.3.2.2 shows that on average there was less than one word difference in how many key-words the children used when talking to KASPAR compared to when they were talking to a human, with the children using approximately 10 key-words in both conditions. The number of key-words the children used with KASPAR ranged from 4 words to 22 words, and with the human interviewer from 2 words to 27 words. With regards to the key-points that the children revealed, we found that the was a median small difference of approximately 1 key-point between the robot and the human conditions, however there was a large range of variation in both conditions, see Table 4.3.2.2.

4.3.1.3. Questionnaire analysis

The questionnaire results suggested that the only considerable difference in how the children evaluated the interviews with KASPAR compared to the human was the average duration of the respective interviews, see Table 4.3.2.3. The children perceived that the interviews with KASPAR were longer, however this perception from the children was correct as can be seen in the results of the interview durations, see Table 4.3.2.1.

4.3.1.4. Crossover effects

Table 4.3.2.4 appears to indicate that there were some crossover effects with regards to the temporal aspects of the interviews, most notably a general tendency for the interviews to be longer in the second phase, and the children to spend slightly more time speaking in the second phase. In addition Table 4.3.2.5 shows that the children had a tendency to use more words in the second phase with a median difference of 41 words. These differences are likely to be due to the children becoming more familiar and comfortable with the procedure.

4.3.2. Data tables

- Table 4.3.2.1 lists the overall temporal measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 4.3.2.2 lists the overall transcription measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 4.3.2.3 lists the basic post-interview questionnaire measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 4.3.2.4 lists the overall temporal measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 4.3.2.5 lists the overall transcription measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 4.3.2.6 lists the basic post-interview questionnaire measurements for both the first and second phase of interviews, and provides statistical data on the differences.

Table 4.3.2.1: The temporal interaction metrics of KASPAR compared to the human interviewer.															
Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interview duration	06:55	06:26	03:44 - 10:45	05:03	05:02	03:24 - 07:33	01:52	01:56	-01:12 - 05:05	01:06	02:38	01:45	00:23	0.105	0.200
Total eye gaze duration	02:23	02:10	01:08 - 06:36	01:24	01:14	00:35 - 03:41	00:58	01:00	-00:33 - 02:55	00:38	01:18	00:46	00:10	0.131	0.200
Proportionate eye gaze duration	0.346	0.322	0.117 - 0.807	0.288	0.273	0.122 - 0.717	0.059	0.066	-0.231 - 0.289	0.009	0.108	0.113	0.025	0.134	0.200
Child response duration	03:53	03:41	01:36 - 07:52	03:21	03:25	01:38 - 06:02	00:32	00:32	-01:55 - 02:54	-00:05	01:10	01:26	00:19	0.094	0.200
Interviewer response duration	00:57	00:57	00:40 - 01:14	00:55	00:55	00:38 - 01:17	00:02	00:00	-00:07 - 00:21	-00:01	00:05	00:07	00:02	0.211	0.019
Response time child > interviewer	01:17	01:18	00:22 - 02:18	00:20	00:19	00:09 - 00:36	00:56	00:57	-00:07 - 01:59	00:44	01:08	00:27	00:06	0.171	0.127
Response time interviewer > child	00:26	00:22	00:09 - 01:01	00:16	00:14	00:05 - 01:08	00:10	00:09	-00:39 - 00:39	00:03	00:17	00:16	00:04	0.171	0.126

Table 4.3.2.2: The transcript interaction metrics of KASPAR compared to the human interviewer.															
Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interviewer word count	153.95	152.5	111 - 214	155.2	157.5	123 - 207	-1.25	-5.5	-41 - 26	-8.952	6.452	17.574	3.930	0.134	0.200
Word count	356.55	363	179 - 672	347.4	329.5	175 - 543	9.15	28.5	-243 - 203	-44.916	63.216	123.364	27.585	0.133	0.200
Proportionate word count	2.342	2.393	0.978 - 4.073	2.267	2.172	1.074 - 3.964	0.074	0.149	-1.655 - 1.346	-0.300	0.449	0.854	0.191	0.181	0.084
Filler word count	18.85	14.5	2 - 101	17.75	13	2 - 62	1.1	0.5	-20 - 39	-4.115	6.315	11.898	2.661	0.160	0.193
Overall key-words	12.25	10	4 - 22	11.9	9.5	2 - 27	0.35	0.5	-7 - 8	-1.170	1.870	3.468	0.776	0.149	0.200
All key points	24.7	25	12 - 48	25.55	21	6 - 69	1.5	-0.85	-36 - 13	-5.839	4.139	11.385	2.546	0.226	0.009

Table 4.3.2.3: The post-interview questionnaire results of the interactions with KASPAR compared to the human interviewer.															
Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
1=boring - 5=interesting	4.7	5	2 - 5	4.55	5	2 - 5	0.15	0	-3 - 3	-0.348	0.648	1.137	0.254	0.298	0.000
1=hard - 5=easy	3.45	4	1 - 5	3.45	4	1 - 5	0	0	-2 - 2	-0.472	0.472	1.076	0.241	0.250	0.002
1=no fun - 5=fun	4.65	5	3 - 5	4.6	5	3 - 5	0.05	0	-1 - 1	-0.174	0.274	0.510	0.114	0.389	0.000
1=long time - 5=quick	3.2	3	1 - 5	4.05	4	1 - 5	-0.85	-0.5	-4 - 2	-1.551	-0.149	1.599	0.357	0.263	0.001

Table 4.3.2.4: The temporal interaction metrics of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interview duration	05:46	05:46	03:24 - 10:00	06:12	05:39	03:58 - 10:45	-00:26	-01:04	-04:20 - 05:05	-01:33	00:41	02:34	00:34	0.134	0.200
Total eye gaze duration	01:39	01:27	00:45 - 03:41	02:07	01:54	00:35 - 06:36	-00:28	-00:35	-02:55 - 01:12	-00:59	00:03	01:10	00:16	0.091	0.200
Proportionate eye gaze duration	0.301	0.309	0.117 - 0.717	0.333	0.314	0.124 - 0.807	-0.032	-0.014	-0.231 - 0.289	-0.086	0.022	0.124	0.028	0.123	0.200
Child response duration	03:26	03:21	01:36 - 06:02	03:48	03:38	01:39 - 07:52	-00:22	-00:19	-02:54 - 02:42	-01:01	00:17	01:29	00:20	0.093	0.200
Interviewer response duration	00:54	00:55	00:38 - 01:14	00:58	00:59	00:40 - 01:17	-00:04	-00:02	-00:21 - 00:03	-00:07	-00:01	00:06	00:01	0.177	0.102
Response time child > interviewer	00:46	00:30	00:16 - 02:18	00:51	00:57	00:09 - 01:54	-00:06	-00:36	-01:38 - 01:59	-00:33	00:22	01:04	00:14	0.233	0.006
Response time interviewer > child	00:22	00:20	00:06 - 01:08	00:20	00:17	00:05 - 01:01	00:02	-00:00	-00:39 - 00:39	-00:06	00:10	00:19	00:04	0.131	0.200

Table 4.3.2.5: The transcript interaction metrics of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interviewer word count	152.4	149.5	119 - 214	157	159	111 - 207	-4.45	0.5	-41 - 23	-11.909	3.009	17.018	3.805	0.211	0.020
Word count	335.5	344	175 - 543	369	352	179 - 672	-33.1	-41	-243 - 169	-85.196	19.096	118.983	26.605	0.078	0.200
Proportionate word count	2.222	2.237	1.074 - 3.964	2.387	2.419	0.978 - 4.073	-0.165	-0.075	-1.646 - 1.655	-0.534	0.203	0.841	0.188	0.119	0.200
Filler word count	16.9	14	2 - 62	19.7	16	2 - 101	-2.8	-2.5	-39 - 20	-7.884	2.284	11.601	2.594	0.159	0.200
Overall key-words	11.95	10	2 - 27	12.2	10	4 - 23	-0.25	1	-8 - 7	-1.774	1.274	3.477	0.778	0.229	0.007
All key points	24.05	21	6 - 48	26.2	23.5	12 - 69	-2.15	0	-36 - 13	-7.060	2.760	11.203	2.505	0.184	0.073

Table 4.3.2.6: The post-interview questionnaire results of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
1=boring - 5=interesting	4.45	5	2 - 5	4.8	5	4 - 5	-0.35	0	-3 - 1	-0.827	0.127	1.089	0.244	0.326	0.000
1=hard - 5=easy	3.4	4	1 - 5	3.5	4	1 - 5	-0.1	0	-2 - 2	-0.569	0.369	1.071	0.240	0.263	0.001
1=no fun - 5=fun	4.55	5	3 - 5	4.7	5	3 - 5	-0.15	0	-1 - 1	-0.364	0.064	0.489	0.109	0.420	0.000
1=long time - 5=quick	3.7	4	1 - 5	3.55	4	1 - 5	0.15	0	-4 - 4	-0.645	0.945	1.814	0.406	0.233	0.006

4.3.3. Data visualisations

The boxplots presents in this section represent the differences between the robotic interviewer and the human interviewer in each of the measures stated.

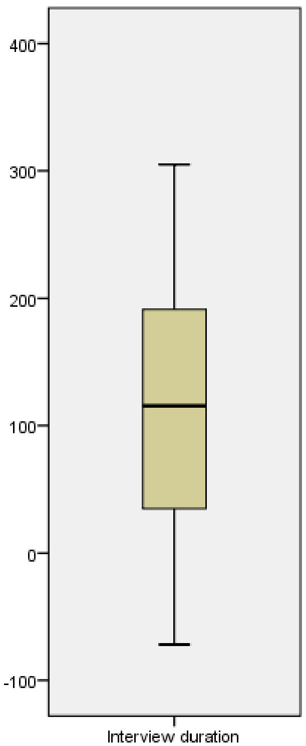


Figure 4.3.3.1: (seconds)

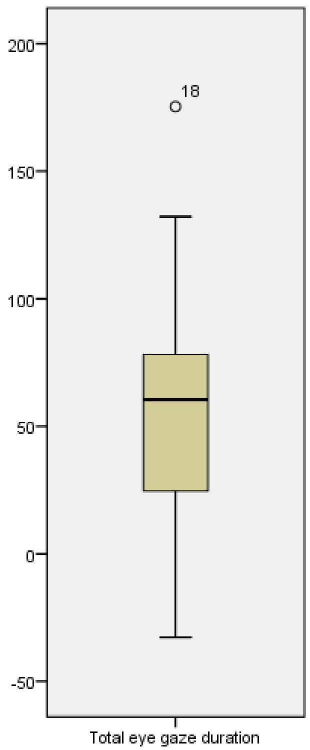


Figure 4.3.3.2: (seconds)

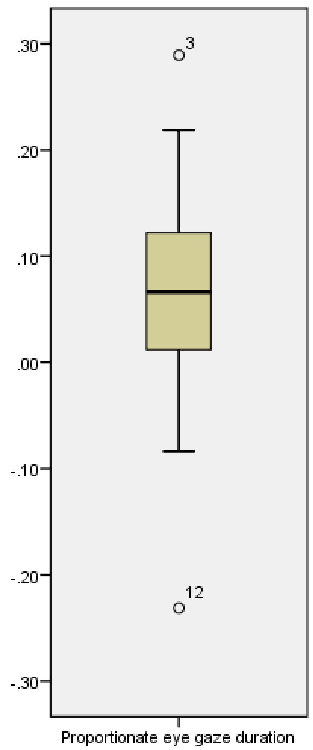


Figure 4.3.3.3: (eye gaze / interview duration)

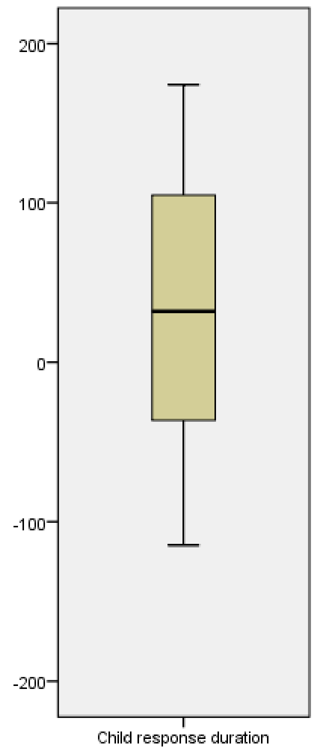


Figure 4.3.3.4: (seconds)

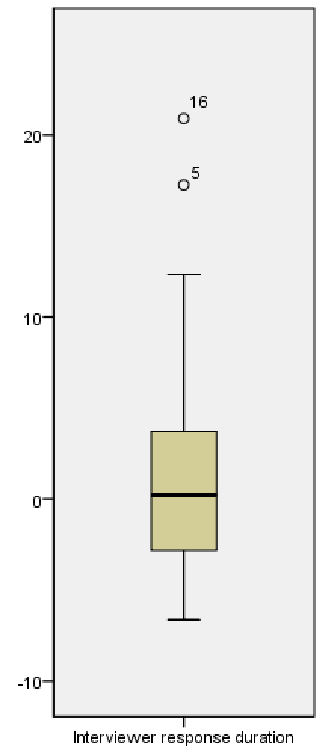


Figure 4.3.3.5: (seconds)

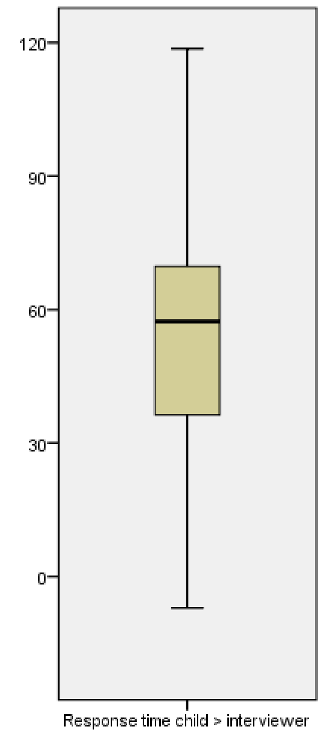


Figure 4.3.3.6: (seconds)

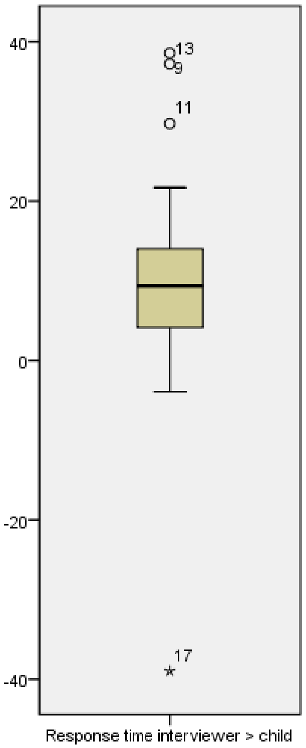


Figure 4.3.3.7: (seconds)

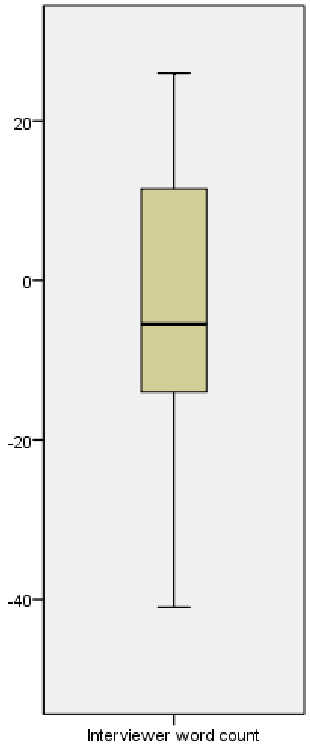


Figure 4.3.3.8: (words)

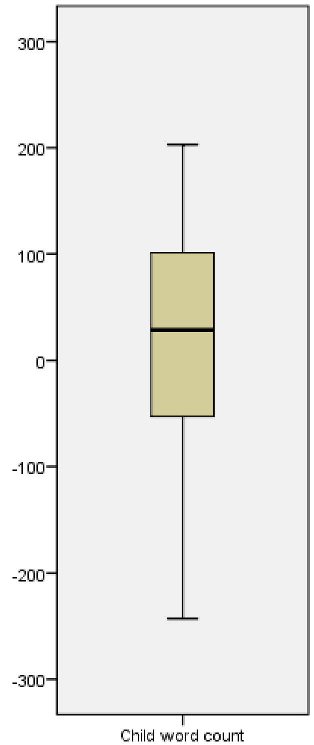


Figure 4.3.3.9: (words)

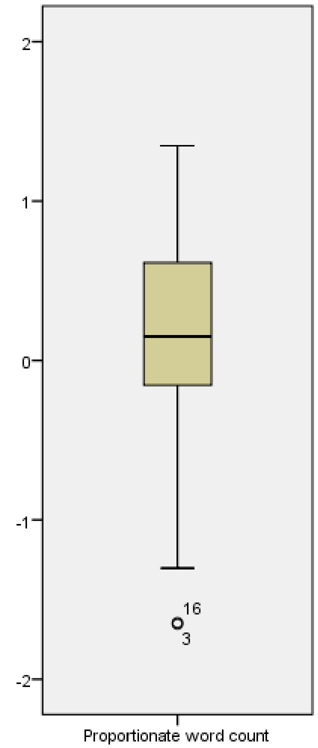


Figure 4.3.3.10: (child word count / interviewer word count)

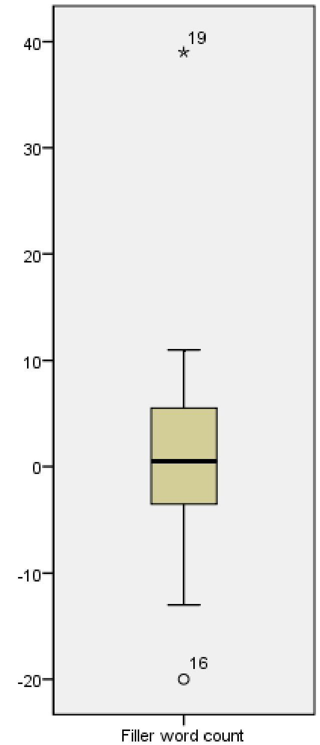


Figure 4.3.3.11: (words)

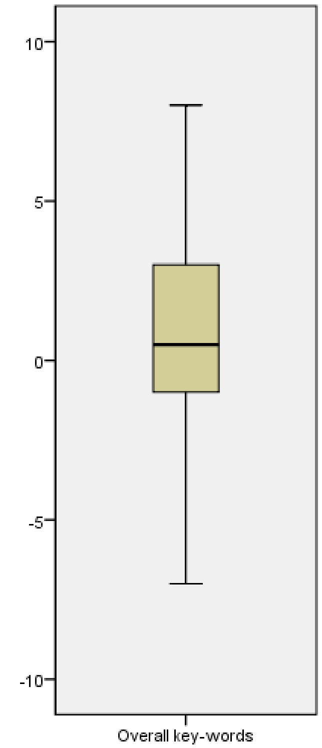


Figure 4.3.3.12: (key-words)

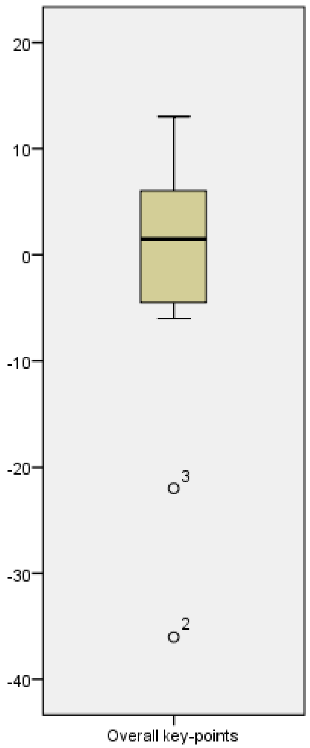


Figure 4.3.3.13: (key-points)

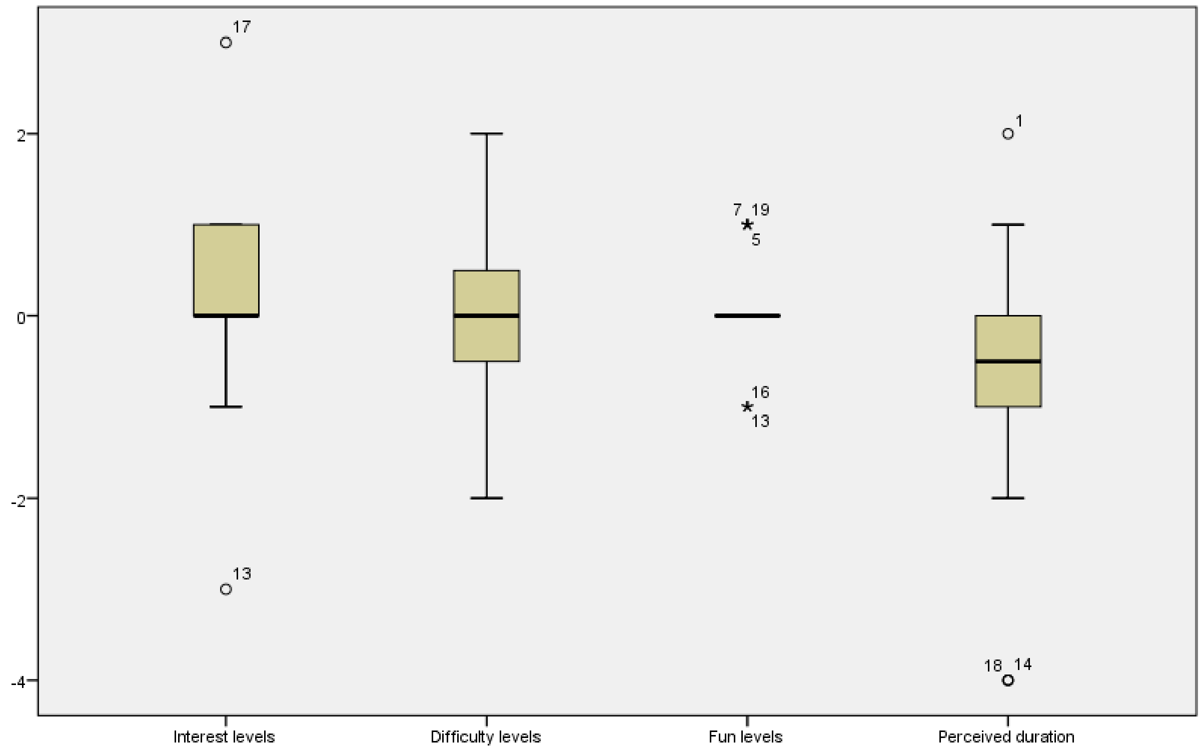


Figure 4.3.3.14: Questionnaire ratings

4.3.4. Reliability of coded data

As stated in Section 3.3.1.4.3, 20% of the video data was coded by an independent reliability coder. The reliability coding in this study was conducted by Chris Miller and a kappa of 0.74 was achieved, which is generally considered good. This result indicates that the coding in this study was conducted to a high standard and the data produced from this coding is therefore reliable [86].

4.4. Discussion

4.4.1. Summary of findings

The results from this study indicate that children were willing to interact with a robot in an interview setting and did so in a similar way to how they interacted with a human interviewer. Furthermore, the amount of information that children provided to KASPAR was also very similar to the information they provided to the human interviewer. This was assessed by measuring the children's use of key-words which we found to be similar in both the robot and human conditions. In addition, the analysis of the key-points again indicated that there was very little difference in the information that the children provided to KASPAR and the human interviewer. There did appear to be some considerable differences in both the duration of the interviews and the eye gaze toward the interviewer. The difference in the duration of the interviews can be explained by the additional time it took for the robot to respond, this was due to the technical limitations of the robot. In the data analysis we found that the robot took much longer to respond to the children and this is why the interviews with the robot took longer, see Table 4.3.2.1. Similarly, the children took longer to respond to the robot, however on average the children only took an additional 9 seconds throughout the entire duration of the interview to respond to the robot, but this may have been because the robot took longer to respond to the children. The delay of the robots response time could have potentially influenced the results of the study if this delay had caused the children to feel a disconnection during the Human-Robot Interaction experience, however this is not supported by the results, possibly because the robot would still blink periodically during the brief periods of delays, thus maintaining the visual appearance of movement and presence of the robot.

In this study there was considerable variation in the durations of the interviews. This was due to the children all being very different in terms of how they spoke and how much information they disclosed. Some children were shy and would not talk much at all whilst others were very confident and would talk for a long time.

The difference in the durations of the interviews was due to the operation of the robot which can be confirmed from the results in Table 4.3.2.1. Getting KASPAR to respond to the children takes longer than it does for a human interviewer present in the room because finding the appropriate key to respond with takes longer, despite extensive training of the operator/experimenter prior to the experiment.

The children's verbal responses were very similar in both conditions with regards to filler words, key-words and key-points. The median difference would suggest that the children used slightly more words with the robotic interviewer (28.5 words). However, the children's word count relative to the interviewers word count was similar. Both interviewers followed the same interview structure and asked the same questions. However, the interviewers are very different in terms of their nature (robot/human), so the similarity in children's responses in both conditions is very encouraging for developing robots as interviewing tools for children. Although the results from this study show that the children interacted with the robot in a similar manner to which they did with a human, and the information they provided was similar, as explained in Chapter 2, there are potential advantages a robot could have over a human interviewer.

The children's similar use of filler words may indicate that the children found talking to KASPAR very similar to talking to the human in terms of comfort. In some respects measuring filler words could provide a better indicator of a child's comfort in a particular situation than a word count. The questions in the interview were focused on an event that took place on one particular day and the interviews were one week apart therefore the amount that the children would remember would inevitably change. The amount of filler words that the children used is likely to be more consistent with the child's level of comfort and the number of questions asked. Some research investigating linguistic disfluencies suggests that the use of filler words could be linked to the difficulty of planning what to say [87, 88]. Whereas other research suggests that filler words may serve a communicative function to help coordinate linguistic interactions [89]. For example, fillers may be used so an individual is not interrupted before they can speak their next sentence [90, 91]. There is also some evidence showing that an increased number of fillers and longer pauses occur before an uncertain answer is given [92, 93]. High disfluency has been associated with anxiety [94]. The children's similar use of filler words in the present experiment may reflect that their comfort levels were the same with both interview partners.

Concerning the effect of the order of the experimental conditions, there were some differences between the first and second rounds of interviews. Table 4.3.2.4 shows that the response durations and response times were generally longed in the second phase. It is likely that the additional time is because over time the children participating in the study and the principle investigator became more comfortable and used to the interview procedure, and as a result took more time responding in the second round of interviews. In addition to this, the children used slightly more words in the second phase, again this is likely due to the children being more familiar with the procedure and as a result were better prepared to answer the questions that they were being asked.

4.4.2. Relation to literature and advances

Generally, the findings from this study are consistent with the HRI literature as the children were happy to talk to and interact with KASPAR. The increased levels of eye gaze also suggest that the children were very interested in KASPAR, however it is possible that the additional eye gaze towards KASPAR occurred because of the additional time it took KASPAR to respond to the children. This study confirms and builds on the findings of the study by Bethel et al. [13] which found that children are equally likely to share a 'secret', or other information, with a robot as they are with a human. The context of the interaction and age ranges differ slightly in the two studies but the basic concept of children talking to a robot is the same. The study conducted as part of this thesis differentiates itself from Bethel et al.'s work in terms of the interview was clearly focused on gathering information from the children and the children would have been aware of this as there were no other tasks to distract them. In addition to this in Bethel et al.'s study the children were told a secret, and then an attempt was made by the interaction partner to obtain the secret from the child. In this study the children were not told to keep a secret, and the children were not pressurised to reveal the correct information. Rather the children were given the freedom to answer the questions as they wished with no emphasis or pressure on honest or correct responses. By giving the children this freedom we found that the children did not need to be pressured in talking to the robot in a similar way to which they did a human interviewer as they would do this naturally, which is a positive indication for the potential of RMI systems.

4.4.3. Limitations

Concerning limitations of this study, the questions used in the interviews were based around a topic of which all the children had very different perceptions. For example some of the children took part in the audition for the event, some took part in the event, two of the children actually won the event as a pair, whilst many of the children only watched the event. This difference in perception would have affected the children's responses although it would not have changed between the interviewers. Another limitation of the study is that the information the children were disclosing was not a 'personal secret' and there was no incentive for the child to keep anything from the interviewer. If the children had a vested interest in keeping information from the interviewer, or if the information had been of a more sensitive nature the results between the human and the robot may have been different. However conducting a study that focuses on questions of a personal matter could be intrusive and would be ethically questionable at such an early stage of research into RMIs.

4.4.4. Review of research questions and hypotheses

This study investigated the difference in how children responded to a robot compared to a human in an interview setting, providing data to answer the first two research questions: "RQ-1: To what degree do children respond to a robot in an interview setting?" and "RQ-2: How do children respond to a robot compared to a human interviewer?". The results show that children looked at KASPAR more than they looked at the human interviewer, which was consistent with the expectations, see Section 4.1.2, possibly because the robot was a novel object to the children and therefore they may have been more interested in KASPAR than the human interviewing them. Ascertaining that children will respond to a robot in an interview setting, as well as to a human is an important first step in this research to establish if robots could be a useful tool for interviewing children.

Contrary to the expectations we found no considerable differences in the amount that the children spoke to KASPAR, the number of key-words the children used with KASPAR, or the amount of key-points the children revealed to KASPAR, compared to the human. We had expected to see a clear preference towards either the robot or the human interviewer, see Section 4.1.2. However, this finding is very encouraging for the future use of robots, as it could be interpreted in such a way that children actually make very little difference between human and robot interviewers in this respect and therefore robot interviewers (i.e. robots as interviewing tools in the hands of experts remotely conducting the interview via the robot) could, with appropriate adjustments, be used as a valuable complement in interviews e.g. with social services, Police and educational psychologists.

The results have shown that children do respond to robots in a similar way in which they respond to a human in an interview setting, providing data to answer the first two research questions and indicating that a RMI approach may have some potential and is worth further investigation.

4.4.5. Research questions for next study

This study provided strong support for continuing the research direction of using robots in an interview setting with young children. The next study investigated how children responded to questions of varying difficulty from both a human and robotic interviewer in an effort to answer the third research question: "RQ-3: Will question type and difficulty impact on children's responses to a robot in an interview setting?".

Chapter 5: Testing the impact of question difficulty on information recovery with a robotic interviewer (Study #2)

5.1. Introduction

Previous studies have explored the possibility of using social robots to recover information from young children [13, 16, 95]. In this chapter we build on and extend the previous study which showed that children interacted with KASPAR very similarly to how they interacted with a human interviewer [95]. However, the previous study did not control for the difficulty of the questions being asked, which may or may not influence how children respond to human or robotic interviewers. When children are being interviewed it is often for a very good reason and some of the questions that they may be asked can be quite difficult for them to answer. Therefore, ascertaining how children respond to different types of question, and more difficult questions, is an important step in establishing if robots could be a useful tool for mediating interviews with young children.

This study investigated how twenty-three children aged between 7 and 9 responded to questions of varying difficulty from a robotic interviewer compared to a human interviewer. Similar to the previous study each child participated in two interviews, one with the human experimenter and one with the humanoid robot KASPAR. The main questions in these interviews focused on the theme of pets and animals. After each interview, the children were asked to rate the difficulty of the questions in addition to the standard measures outlined in Section 3.3. The results from quantitative data analysis reveal that the children interacted with KASPAR in a very similar manner to how they interacted with the human interviewer. Similar to the first study the children provided both interviewers with comparable information and amounts of information regardless of question difficulty.

5.1.1. Research questions

The first study in Chapter 4 has shown that children respond to a robotic interviewer in a very similar way to which they do a human interviewer, answering the first two research questions in this thesis: "RQ-1: To what degree do children respond to a robot in an interview setting?" and "RQ-2: How do children respond to a robot compared to a human interviewer?". The positive results from the first study encouraged us to continue investigating the possibility of RMIs further. The second follow-up study in this chapter aimed to verify the results of the first study with regards to the first two research questions, and build on this by answering the third research question: "RQ-3: Will question type and difficulty impact on children's responses to a robot in an interview setting?".

5.1.2. Hypothesis

The hypothesis in this study was that the children would reveal either an equal amount of or more information to a robot than a human interviewer. The reasoning for this is that the children in the previous study responded to the robot in a similar manner to which they did with a human interviewer, sharing similar amounts of information. We therefore hypothesized that the children's performance with the robot would at least be equal or possibly better with the robot than the human interviewer when presented with more difficult questions. The children may be more relaxed and feel less pressured with the robot than the human interviewer because they may view the robot as more of a peer than the adult interviewer.

5.2. Methods

5.2.1. Participants

This study was conducted in a mainstream primary school in Hertfordshire (UK) with twenty-three children, twenty (8 male, 12 female) of which produced useable data[5]. The children were aged between 7 and 9 with an average age of 8 years 10 months, none of the children had interacted with KASPAR before. In accordance with the ethics procedure, informed consent was obtained in writing from the parents of all the children participating in the study as outlined in Section 3.3.1.4.3. The ethics protocol number for this study is 1112/167, the relevant documentation can be seen in Appendix 3.1.

5.2.2. Procedure

The procedure used in this experiment was identical to the first experiment with the children receiving a group briefing before the study commenced, two interviews and a group debriefing at the end of the study where the children would be shown how the robot works and given the opportunity to control the robot. The only difference with the procedure in this study was an additional questionnaire at the end of the study where the children would rate the difficulty of the questions that they had been asked during the interview.

[5] Due to technical difficulties in one interview and non-attendance of the second interview by two of the children, three of the sessions were not included.

5.2.3. Experimental setup

The interviews were conducted in an unused classroom that contained a small lockable cupboard, which was used as a control room for KASPAR. The control room housed a monitor with a wireless connection to camera #1, which was used to observe the situation and make KASPAR respond appropriately, and a laptop with a remote connection to the robot. Figure 5.2.3.1 show the room layout, which was arranged to an appropriate specification as outlined in Section 3.2.8. The interviews were always led by the principle investigator either in person or remotely via KASPAR to maintain consistency between the interviews. The children were unaware that KASPAR was being controlled by a human triggering the correct questions and responses from a pre-recorded list. In the first study, the principle investigator found it difficult to locate the keys quickly on the keyboard from the reference sheet. In order to assist with locating keys quickly the keys on the keyboard were labeled with colored and numbered labels in this study. By using colours and numbers it was easier to locate the appropriate keys quickly, which assisted in controlling the robot, see Figure 5.2.3.1. A second research assistant unknown to the children took the children to and from the interviews and remained in the room during the interviews, but was as non-reactive as possible. Immediately after each interview the children were asked by the research assistant to rate the difficulty of the questions that they had just been asked along with other details about the interview.

Figure 5.2.3.1: An overview of the room layout and interview setup

5.2.4. The interview

The interviews were conducted on three days over a two-week period using the two-phase counterbalancing method described in Section 3.2.7 with each child experiencing two interviews one week apart[6]. The interviews began with a short introduction and some simple rapport building questions that would establish the child's name, age and whether they have any siblings. This was followed by a question asking the children if they had, or would like pets, in preparation for the main topic of pets and animals. The majority of the questions were open questions or were followed by a descriptive question to encourage the children to elaborate on the details to maximize their freedom to express themselves. Research and practice indicates that the most detailed and reliable answers are secured from open questions [18],p.27. The interviews concluded by thanking the child for their time and participation. As stated in Section 2.1.1, the structure and questions for the interviews were derived from the ABE guidance document used by the UK Police, as this is a recognized standard approach to interview children [18].

The main interview questions were specifically designed to vary in difficulty because we sought to investigate if the children's responses to a robot would differ to their responses with a human interviewer when faced with questions of varying difficulty. To ensure that the questions were appropriate and at a suitable level for the children, their class teacher rated a selection of questions prior to the study commencing. From the list of questions ten were selected to be used in the interviews, as they represented a range of levels of difficulty. In addition, the children were asked to rate the difficulty of the questions at the end of the interview. This was particularly useful because each child is different. For example, if one child had recently visited a zoo it is likely that they may find some questions easier than another child that has not visited the zoo. Also, a child that is only just 8 years old may find some questions more difficult than a child that is 9 and a half years old.

In this study the main interview questions were designed to recover the children's knowledge, opinion and experience of pets and animals. In addition to this we also recovered basic information about the child's own pets that was personally relevant to them and the subject of the interview, but not of a sensitive nature. Having a range of different types of question allowed us to investigate if there is any difference in how children respond to a robot compared to a human when faced with different types of question as well as question difficulty. See Table 5.3.2.1 for the information categories and question types.

[6] There were five exceptions due to late consent form submission, these children had their interviews one day apart. The results from these sessions were consistent with the data from the rest of the study and were therefore included in the final dataset.

5.2.5. Measurements

The primary sources of data for this study were derived from the standardized measurements outlined in Section 3.3. In addition to these measures an extra questionnaire was created to capture the children's perceptions of the question difficulties. Immediately after each interview the child was asked to rate the difficulty of the ten main interview questions as easy, medium or hard to answer. The difficulty measure allowed us to assess if the question difficulty was affected by the interviewer, see Figure 5.2.5.1.

Questions	Difficulty		
	Easy	Medium	Hard
Do you have any pets?			
Why do people have pets?			
What can children learn by having a pet?			
Can you tell me what a zoo is?			
What animals would you see at a zoo?			
What animal do you think is the most dangerous?			
What is the strangest animal you have seen?			
What is the largest animal you have seen?			
Can you name some animals that live in this country?			
If you were an animal what animal would you like to be?			

Figure 5.2.5.1: Sample of question difficulty questionnaire

5.2.5.1. Communicative content

In this study we conducted a detailed analysis of the transcriptions, examining and logging the details (key-points) of what was said in response to the questions the children were asked. A key-point was defined as a specific piece of information that we recovered from the child relating to the question they had been asked. When analysing the key-points we compared the points given in both experimental conditions, to establish the consistency of the facts disclosed in both interview conditions. This information consisted of 11 main categories that cover the following areas:

- **Relatives** - Since one of the interview questions asked about the children's siblings, in this category we analysed how many family members the children mention in total throughout the duration of the interview, and how many family members they state by name. Although this was an introductory rapport building question and not the primary focus of the interview, it is something that relates directly to the children and could be an indicator as to how much they are willing to talk about their family.

- **Pets** - The interview question about pets served as both an introductory rapport building question and related to the primary topic of the interview. Similar to the relative's category, we analysed how many pets the children mention and how many pets they state by name.

- **Reasons people have pets** - The children were asked "Why do people have pets?". For this category each reason that the child stated was listed as a key-point. A reason was not necessarily a particular word it was often several words i.e. to look after, to play with.

- **What children learn from pets** - The children were asked "What can children learn by having a pet?". For this category we listed all of the things that the children thought a child could learn by having a pet i.e. how to care for them, sometimes they can bite.

- **What a zoo is** - The children were asked "Can you tell me what a zoo is?". For this category we listed the points that the children used to describe a zoo i.e. Keep animals, Look at animals, Wild animals.

- **Zoo animals** - Following the question about the zoo the children were asked "What animals would you see at a zoo". For this question each animal that the child listed was noted as a key-point and any duplications were not included when noting the key-points.

- **Most dangerous animal and why** - After listing animals that you would see at a zoo the children were asked "What animal do you think is the most dangerous". The key-points that were logged from the children's responses were the animal(s) the child considered to be the most dangerous and the reasons why i.e. Lion, sharp-teeth, eat-meat.

- **Strangest animal and what it looks like** - After the question about the most dangerous animals the children were asked "What is the strangest animal that you have seen". The key-points logged from this question were the animal(s) that the children stated and the description the children gave of the animal i.e. Otter, slimy, fast in water.

- **Largest animal and what it looks like** - Following the question about the strangest animal the children were asked "What is the largest animal you have seen". The key-points we took from this was the animal(s) and the details of the description i.e. Giraffe, long neck, small face, tall legs.

- **Animals that live in this country** - Towards the end of the interview the children were asked "Can you name some animals that live in this country". For this question each animal that the child listed was noted as a key-point and any duplicates were not included.
- **Animal you would like to be and why** - The last question that the children were asked was "If you were an animal what animal would you like to be". For this question we noted the animal(s) that the child would choose to be and the reasons why. i.e. Cat, very quick, quite cute.

5.3. Results

5.3.1. Data analysis

Similar to the first study we processed and tested the data as outlined in Section 3.4, to ascertain the similarities and difference between the human and the robotic interviewer conditions. In this study a much more detailed analysis of the responses to the individual questions was performed as we were analysing the impact of question difficulty on the children's responses to the two interviewers.

5.3.1.1. Temporal analysis

Examining the interview durations between the human and robotic interviewers we found that the interviews with KASPAR lasted approximately 20 to 25 seconds longer, see Table 5.3.2.3. On average the interviews with KASPAR lasted approximately 6 minutes, and interviews with the human lasted approximately 5 minutes 40 seconds, although there was considerable variation in the durations of the interviews, see Figure 5.3.3.1. Interviews with KASPAR ranged from 4:26 to 9:26, whilst interviews with the human experimenter were between 2:18 to 8:35. Assessing the different response durations, pause durations and response times for both the child and the interviewer, we found that similar to the first study, there appears to be a large differences in terms of the interviewers response time. KASPAR took much longer to respond to the children than the human interviewer due to the technical limitations of the system. Throughout the full duration of the interviews KASPAR took approximately 1:27 to respond to the children, while the human interviewer took 21 seconds, however there were some extreme cases that fell above and below this, see Figure 5.3.3.4. There also appears to be a slight difference in the time the interviewer spent responding to the children, with the human interviewer spending more time responding to the children. On average KASPAR spoke for 52 seconds whilst the human interviewer spoke for 1 minute 5 seconds. There was only a very slight difference in the amount of time the children spent speaking to each interviewer, with a median difference of 6 seconds less time speaking to KASPASR than the human interviewer, see Table 5.3.2.3. However as with the previous study there was a lot of variation in the amount of time that the children would spend speaking to each interviewer, see Figure 5.3.3.2. There was a negligible average differences in the amount of time the children spent responding to each interviewer and the children's pause duration, however as with the children's speaking duration, there was a vast amount of variation between the different children, see Figure 5.3.3.2. Contrary to the first study there was very little differences in the amount of eye gaze towards the different interviewers. Although the median total eye gaze towards KASPAR was 22 seconds longer than with the human interviewer throughout the full duration of the interviews (see Figure 5.3.3.7), proportionate to the interview duration these averages was almost no difference, see Figure 5.3.3.8.

5.3.1.2. Textual content analysis

The amount of words that the children spoke to each interviewer did vary greatly as can be seen in Figure 5.3.3.9. However, there was an overall median difference of just 6 words between the interviewers in terms of how many words the children used throughout the entire interview, see Table 5.3.2.4. The number of filler words that the children used was similar in both conditions, on average the children used 9 filler words with KASPAR and 10 words with the human interviewer. Analysing the filler word count proportionality also revealed that there is very little difference in the amount of filler words the children use with KASPAR compared to the human interviewer, see Figure 5.3.3.12. There was however a considerable difference in the amount of words that the interviewers spoke to the children, see Figure 5.3.3.13. On average KASPAR spoke 150 words to the children and the human interviewer spoke 180 words, see Table 5.3.2.4. With regards to key-words, there was a small median difference of 1.5 words in the amount of key-words that the children used with KASPAR compared to the human interviewer, with the children using slightly less with KASPAR, see Figure 5.3.3.14. However proportionately the number of key-words the children used with KASPAR was slightly higher overall with a median of 0.153 with KASPAR and 0.125 with the human interviewer, see Figure 5.3.3.15. The results of the key-point analysis reveal that the interviewer did have an effect on the amount of key-points the children revealed in the interviews, with a median difference of 2.5 suggesting that the children reveal slightly more key points to the human interviewer overall, see Figure 5.3.3.16.

5.3.1.3. Basic questionnaire analysis

The experience questionnaire results suggested that there were some differences in how interesting the children found the experience and how difficult they found the experience. The children found talking to KASPAR more interesting, but the children also found talking to KASPAR more difficult, see

Figure 5.3.3.17. The median values calculated suggest that there was no real difference in how fun they found the activity or how long they perceived the interviews took between each interviewer, however Figure 5.3.3.17 suggests that there was some variation between the children.

5.3.1.4. Question difficulty ranking

The design of the experiment required different difficulty questions in order to establish if there was a difference between how children responded to a robot compared to a human interviewer. The results from the questionnaires show that we were successful in the design of this experiment, with a set of questions that the children felt ranged from easy to difficult. To calculate the ranges and boarders for the question difficulties we took the highest average difficulty 2.45 and subtracted the lowest average difficulty 1.05 to acquire the difference 1.40. We then divided this average difference by 3 to set the difficulty boundaries which were as follows: Easy=1.05-1.52, Medium=1.52-1.98, Hard=1.98-2.45. Although there is some variability in the ranking of the questions between interviewers and phase, this is likely to be due to different children having different personal opinions. Table 5.3.2.18 shows that on average there was very little difference between how the children scored the question difficulties between the two interviewers and that the median difference was zero for each category. Using the averages to calculate, there were 4 easy questions, 3 medium difficulty questions, and 3 hard questions, see Table 5.3.2.2.

5.3.1.5. Question specific data

In addition to the ranking data from the questionnaires, the specific word counts and child speaking durations for each question were analysed. When comparing the differences, the median durations would suggest that the children generally took slightly longer talking to KASPAR for most of the questions, with the exception of one which the mean durations were equal for and 2 which the children appear to talk to the human interviewer for longer, see Figure 5.3.3.20. Investigating this further, the child word counts for many of the each of the main questions appears to be slightly lower for KASPAR, see Figure 5.3.3.19. Giving the question durations further consideration, in comparison to all of the other questions, the two questions that required the children to name a list from their memory were the questions that they spent more time talking to the human interviewer for. No firm conclusions can be drawn from this one result, however it may be that children do not perform listing activities as well with a robot as they do with a human interviewer.

5.3.1.6. Crossover effects

As can be seen from Tables 5.3.2.9 to 5.3.2.14, the majority of results were relatively consistent between the first and second phase of the interviews. In contrast to the first study, the word counts of the children were very similar to the first phase, with a mean difference of just 6.5. Table 5.3.2.10 shows that there is a notable difference between the amount of key-words and key-points that the children use, with the children revealing more in the second interview, however this is likely to be because they were more familiar with the procedure and had recently been asked the same questions. Table 5.3.2.9 shows that the children also felt that the interviews took less time second time round with a median difference of 1.

5.3.2. Data tables

- Table 5.3.2.1 provides a list of the main interview questions and the category in which they fall into.
- Table 5.3.2.2 illustrates overall how the children rated the questions in terms of difficulty between the different interviewers, different phases and the overall difficulty for each question.
- Table 5.3.2.3 lists the overall temporal measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.4 lists the overall transcription measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.5 lists the basic post-interview questionnaire measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.6 lists the detailed post-interview question difficulty ratings for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.7 lists the question specific transcription measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.8 lists the question specific temporal measurements for both the robotic and human interviewer, and provides statistical data on the differences.
- Table 5.3.2.9 lists the overall temporal measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 5.3.2.10 lists the overall transcription measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 5.3.2.11 lists the basic post-interview questionnaire measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 5.3.2.12 lists the detailed post-interview question difficulty ratings for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 5.3.2.13 lists the question specific transcription measurements for both the first and second phase of interviews, and provides statistical data on the differences.
- Table 5.3.2.14 lists the question specific temporal measurements for both the first and second phase of interviews, and provides statistical data on the differences.

Table 5.3.2.1: The ten main interview questions and corresponding categories				
Question	**Personal**	**Knowledge**	**Opinion**	**Experience**
Do you have any pets?	x			x
Why do people have pets?			x	
What can children learn by having a pet?			x	x
Can you tell me what a zoo is?		x		x
What animals would you see at a zoo?		x		
What animal do you think is the most dangerous, and why?		x	x	
What is the strangest animal you have seen, and what it looks like?		x	x	x
What is the largest animal you have seen, and what it looks like?		x		x
Can you name some animals that live in this country?		x		
If you were an animal what animal would you like to be, and why?			x	

Table 5.3.2.2: Question difficulty as rated by the children					
Question	**KASPAR**	**Human**	**Phase 1**	**Phase 2**	**Average**
Do you have any pets?	Easy	Easy	Easy	Easy	Easy
Why do people have pets?	Hard	Hard	Hard	Hard	Hard
What can children learn by having a pet?	Hard	Hard	Hard	Hard	Hard
Can you tell me what a zoo is?	Easy	Easy	Medium	Easy	Easy
What animals would you see at a zoo?	Easy	Medium	Medium	Easy	Medium
What animal do you think is the most dangerous?	Medium	Medium	Medium	Medium	Medium
What is the strangest animal you have seen?	Hard	Medium	Hard	Medium	Medium
What is the largest animal you have seen?	Medium	Easy	Easy	Medium	Easy
Can you name some animals that live in this country?	Hard	Hard	Hard	Hard	Hard
If you were an animal what animal would you like to be?	Easy	Easy	Easy	Easy	Easy

Table 5.3.2.3: The temporal interaction metrics of KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interview duration	06:00	05:56	04:26 - 09:26	05:39	05:44	02:18 - 08:35	00:21	00:23	-02:39 - 02:52	00:11	00:31	01:31	00:05	0.103	0.200
Child response duration	02:19	02:29	00:50 - 05:30	02:47	02:51	00:35 - 05:43	-00:28	-00:06	-02:57 - 01:22	-00:25	-00:30	01:06	-00:01	0.207	0.024
Interviewer response duration	00:52	00:53	00:47 - 00:56	01:06	01:05	00:54 - 01:20	-00:14	-00:13	-00:29 - -00:01	-00:08	-00:20	00:07	-00:03	0.131	0.200
Response time child > interviewer	01:29	01:27	00:53 - 02:20	00:28	00:21	00:10 - 01:38	01:01	01:06	-00:29 - 01:44	00:32	01:30	00:30	00:15	0.247	0.002
Response time interviewer > child	00:33	00:32	00:17 - 00:55	00:33	00:24	00:13 - 01:30	-00:00	00:03	-00:57 - 00:39	-00:02	00:01	00:20	00:01	0.192	0.052
Child pause duration	00:37	00:37	00:06 - 01:17	00:39	00:35	00:03 - 01:38	-00:02	-00:05	-00:47 - 00:44	00:00	-00:04	00:26	-00:01	0.124	0.200
Total eye gaze duration	01:51	01:44	00:50 - 03:01	01:38	01:35	00:39 - 02:54	00:12	00:22	-01:09 - 01:06	00:03	00:22	00:35	00:05	0.149	0.200
Proportionate eye gaze duration	0.319	0.331	0.141 - 0.558	0.311	0.280	0.076 - 0.589	0.008	0.000	-0.259 - 0.246	0.008	0.008	0.133	0.000	0.099	0.200

Table 5.3.2.4: The transcript interaction metrics of KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Child word count	241.3	231.5	65 - 710	308.05	277.5	54 - 641	-66.75	-5.5	-370 - 254	-64.340	-69.160	146.844	-1.230	0.198	0.039
Proportionate child word count	1.615	1.598	0.417 - 4.765	1.737	1.538	0.290 - 3.684	-0.122	0.120	-1.877 - 1.950	-0.175	-0.070	0.873	0.027	0.178	0.095
Filler word count	8.55	6	0 - 22	9.75	7.5	0 - 33	-1.2	-1	-15 - 11	-0.762	-1.638	6.271	-0.224	0.213	0.018
Proportionate Filler word count	0.038	0.034	0 - 0.089	0.034	0.027	0 - 0.079	0.004	0.006	-0.034 - 0.054	0.002	0.007	0.019	0.001	0.124	0.200
Interviewer word count	150.4	151	137 - 160	180	180	162 - 206	-29.6	-30.5	-62 - -7	-16.233	-42.967	11.749	-6.820	0.173	0.119
Overall key-words	33.7	35	14 - 58	37.4	39.5	12 - 64	-3.7	-1.5	-25 - 15	-3.043	-4.357	11.216	-0.335	0.162	0.176
Proportionate key-word	0.159	0.153	0.082 - 0.247	0.137	0.125	0.08 - 0.22	0.022	0.024	-0.058 - 0.105	0.011	0.032	0.043	0.005	0.100	0.200
Overall key-points	32.6	31.5	18 - 44	35.9	38.5	14 - 52	-3.3	-2.5	-16 - 9	-2.204	-4.396	7.760	-0.559	0.165	0.159

Table 5.3.2.5: The post-interview questionnaire results of the interactions with KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
1=boring - 5=interesting	4.175	4	3 - 5	3.4	3	1 - 5	0.775	0.75	0 - 4	0.446	1.104	1.006	0.168	0.262	0.001
1=hard - 5=easy	3.4	3.5	1 - 5	4.075	4.5	2 - 5	-0.675	-1	-4 - 2	-0.237	-1.113	1.321	-0.224	0.203	0.030
1=no fun - 5=fun	3.8	4	1 - 5	3.6	4	1 - 5	0.2	0	-2 - 3	0.200	0.200	1.005	0.000	0.279	0.000
1=long time - 5=quick	3.65	4	1 - 5	3.65	4	1 - 5	0	0	-2 - 4	0.000	0.000	1.414	0.000	0.190	0.057

Table 5.3.2.6: The post-interview question difficulty questionnaire results of the interactions with KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Average question difficulty	1.715	1.65	1.3 - 2.8	1.74	1.8	1.3 - 2.7	-0.025	0	-0.6 - 0.5	-0.025	-0.025	0.243	0.000	0.159	0.200
Do you have any pets?	1.2	1	1 - 2	1.05	1	1 - 2	0.15	0	0 - 1	0.150	0.150	0.366	0.000	0.509	0.000
Why do people have pets?	2.15	2	1 - 3	2.3	2	1 - 3	-0.15	0	-1 - 1	-0.150	-0.150	0.671	0.000	0.288	0.000
What can children learn by having a pet?	2	2	1 - 3	2.05	2	1 - 3	-0.05	0	-1 - 1	-0.050	-0.050	0.510	0.000	0.389	0.000
Can you tell me what a zoo is?	1.45	1	1 - 3	1.5	1	1 - 3	-0.05	0	-1 - 1	-0.050	-0.050	0.510	0.000	0.389	0.000
What animals would you see at a zoo?	1.5	1	1 - 3	1.65	2	1 - 2	-0.15	0	-1 - 1	-0.150	-0.150	0.587	0.000	0.351	0.000
What animal do you think is the most dangerous?	1.7	1.5	1 - 3	1.9	2	1 - 3	-0.2	0	-2 - 1	-0.200	-0.200	1.056	0.000	0.225	0.009
What is the strangest animal you have seen?	2	2	1 - 3	1.95	2	1 - 3	0.05	0	-2 - 1	0.050	0.050	0.887	0.000	0.228	0.008
What is the largest animal you have seen?	1.6	1	1 - 3	1.35	1	1 - 3	0.25	0	-1 - 2	0.250	0.250	0.716	0.000	0.336	0.000
Can you name some animals that live in this country?	2.2	2	1 - 3	2.2	2	1 - 3	0	0	-2 - 1	0.000	0.000	0.649	0.000	0.400	0.000
If you were an animal what animal would you like to be?	1.35	1	1 - 3	1.45	1	1 - 3	-0.1	0	-1 - 1	-0.100	-0.100	0.553	0.000	0.372	0.000

Table 5.3.2.7: The question specific transcript interaction metrics of KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Do you have any pets?	23.9	11	0 - 190	19.15	11.5	1 - 68	4.75	1	-68 - 127	4.312	5.188	36.456	0.224	0.250	0.002
Why do people have pets?	15.2	12	3 - 62	18.35	11.5	2 - 46	-3.15	-0.5	-22 - 25	-2.931	-3.369	12.700	-0.112	0.172	0.124
What can children learn by having a pet?	15.95	10.5	2 - 69	19.8	16.5	0 - 50	-3.85	-0.5	-28 - 33	-3.631	-4.069	15.226	-0.112	0.174	0.113
Can you tell me what a zoo is?	21.55	18.5	1 - 70	27.85	24.5	0 - 85	-6.3	0.5	-56 - 19	-6.519	-6.081	17.257	0.112	0.192	0.050
What animals would you see at a zoo?	14.1	11.5	5 - 34	23.15	20.5	5 - 62	-9.05	-3.5	-43 - 6	-7.516	-10.584	14.248	-0.783	0.235	0.005
What animal do you think is the most dangerous?	22.15	20.5	4 - 53	27.4	21	4 - 78	-5.25	-5	-43 - 25	-3.059	-7.441	16.029	-1.118	0.158	0.200
What is the strangest animal you have seen?	22.9	20.5	4 - 50	31.8	24.5	2 - 110	-8.9	-2.5	-78 - 26	-7.804	-9.996	24.535	-0.559	0.197	0.041
What is the largest animal you have seen?	9.75	2.5	1 - 54	13.8	3	1 - 55	-4.05	0	-51 - 27	-4.050	-4.050	18.802	0.000	0.393	0.000
Can you name some animals that live in this country?	12.65	9	3 - 46	17.65	15	0 - 72	-5	0	-63 - 23	-5.000	-5.000	17.363	0.000	0.223	0.010
If you were an animal what animal would you like to be?	24.3	21.5	3 - 54	33.05	25.5	5 - 142	-8.75	-3	-125 - 20	-7.435	-10.065	31.041	-0.671	0.321	0.000

Table 5.3.2.8: The question specific temporal interaction metrics of KASPAR compared to the human interviewer.

Measure	KASPAR			Human			Difference (KASPAR – Human)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Do you have any pets?	00:24	00:21	00:00 - 01:38	00:17	00:14	00:03 - 00:45	00:07	00:05	-00:45 - 00:58	00:05	00:09	00:23	00:01	0.217	0.014
Why do people have pets?	00:21	00:16	00:10 - 00:45	00:22	00:20	00:06 - 00:48	-00:01	00:00	-00:32 - 00:27	-00:01	-00:01	00:15	00:00	0.136	0.200
What can children learn by having a pet?	00:20	00:16	00:09 - 00:43	00:18	00:16	00:05 - 00:32	00:01	00:02	-00:14 - 00:14	00:00	00:02	00:08	00:00	0.081	0.200
Can you tell me what a zoo is?	00:20	00:21	00:09 - 00:41	00:19	00:16	00:07 - 00:45	00:01	00:02	-00:21 - 00:15	00:00	00:02	00:09	00:00	0.115	0.200
What animals would you see at a zoo?	00:31	00:30	00:18 - 00:47	00:37	00:35	00:18 - 01:35	-00:06	-00:07	-00:48 - 00:12	-00:04	-00:09	00:14	-00:01	0.198	0.039
What animal do you think is the most dangerous?	00:30	00:29	00:14 - 00:54	00:30	00:30	00:12 - 00:49	-00:00	00:02	-00:21 - 00:26	-00:01	00:01	00:12	00:00	0.129	0.200
What is the strangest animal you have seen?	00:37	00:39	00:16 - 01:05	00:34	00:31	00:15 - 01:46	00:03	00:04	-01:02 - 00:50	00:01	00:05	00:21	00:01	0.181	0.085
What is the largest animal you have seen?	00:19	00:14	00:09 - 00:45	00:19	00:11	00:06 - 00:51	-00:00	00:03	-00:32 - 00:25	-00:01	00:01	00:14	00:01	0.253	0.002
Can you name some animals that live in this country?	00:34	00:34	00:13 - 01:01	00:36	00:37	00:10 - 01:12	-00:02	-00:06	-00:40 - 00:38	00:01	-00:04	00:20	-00:01	0.132	0.200
If you were an animal what animal would you like to be?	00:30	00:27	00:17 - 00:49	00:26	00:24	00:11 - 01:20	00:04	00:06	-01:01 - 00:24	00:01	00:07	00:18	00:01	0.227	0.008

Table 5.3.2.9: The temporal interaction metrics of phase 1 compared to phase 2.															
Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Interview duration	05:49	05:49	03:21 - 09:26	05:50	05:59	02:18 - 08:35	-00:00	00:02	-02:39 - 02:52	-00:02	00:01	01:34	00:01	0.065	0.200
Child response duration	02:30	02:33	00:35 - 05:30	02:36	02:28	00:41 - 05:43	-00:06	-00:08	-02:57 - 02:34	-00:02	-00:09	01:11	-00:02	0.164	0.167
Interviewer response duration	01:01	01:01	00:47 - 01:20	00:58	00:55	00:49 - 01:12	00:03	00:09	-00:25 - 00:29	-00:01	00:08	00:15	00:02	0.180	0.088
Response time child > interviewer	00:58	01:02	00:14 - 01:50	01:00	01:08	00:10 - 02:20	-00:02	00:08	-01:44 - 01:30	-00:06	00:01	01:09	00:02	0.204	0.029
Response time interviewer > child	00:36	00:28	00:15 - 01:30	00:29	00:26	00:13 - 00:56	00:07	00:02	-00:17 - 00:57	00:06	00:08	00:19	00:00	0.192	0.052
Child pause duration	00:35	00:30	00:03 - 01:17	00:41	00:43	00:06 - 01:38	-00:06	00:05	-00:47 - 00:28	-00:08	-00:04	00:25	00:01	0.217	0.015
Total eye gaze duration	01:44	01:38	00:50 - 03:01	01:46	01:41	00:39 - 02:54	-00:02	-00:03	-01:09 - 01:06	-00:00	-00:03	00:37	-00:01	0.117	0.200
Proportionate eye gaze duration	0.305	0.300	0.141 - 0.490	0.325	0.315	0.076 - 0.589	-0.020	-0.033	-0.259 - 0.234	-0.006	-0.035	0.131	-0.007	0.110	0.200

Table 5.3.2.10: The transcript interaction metrics of phase 1 compared to phase 2.															
Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Child word count	274.85	251.5	54 - 710	274.5	266	65 - 641	0.35	-6.5	-370 - 345	3.199	-2.499	162.028	-1.453	0.204	0.028
Proportionate child word count	1.655	1.560	0.290 - 4.765	1.696	1.609	0.417 - 3.684	-0.041	-0.120	-1.877 - 1.950	0.012	-0.093	0.881	-0.027	0.140	0.200
Filler word count	9.45	8	0 - 21	8.85	5.5	0 - 33	0.6	0.5	-15 - 12	0.381	0.819	6.361	0.112	0.141	0.200
Proportionate Filler word count	0.038	0.037	0 - 0.079	0.033	0.025	0 - 0.089	0.005	0.005	-0.025 - 0.054	0.003	0.007	0.019	0.001	0.160	0.191
Interviewer word count	170.5	175.5	137 - 206	159.9	156.5	139 - 188	10.6	27.5	-32 - 62	-1.452	22.652	30.693	6.149	0.236	0.005
Overall key-words	34.3	34.5	12 - 61	36.8	36.5	12 - 64	-2.5	-3.5	-25 - 25	-0.966	-4.034	11.560	-0.783	0.117	0.200
Proportionate key-word	0.144	0.146	0.08 - 0.22	0.152	0.151	0.09 - 0.247	-0.008	-0.007	-0.105 - 0.083	-0.004	-0.011	0.048	-0.002	0.081	0.200
Overall key-points	32.85	34	14 - 51	35.65	38	14 - 52	-2.8	-2.5	-16 - 13	-1.704	-3.896	7.964	-0.559	0.093	0.200

Table 5.3.2.11: The post-interview questionnaire results of phase 1 compared to phase 2.															
Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
1=boring - 5=interesting	3.65	4	1 - 5	3.925	4	2 - 5	-0.275	0	-4 - 2	-0.275	-0.275	1.251	0.000	0.263	0.001
1=hard - 5=easy	3.725	3.75	2 - 5	3.75	4	1 - 5	-0.025	0	-2 - 4	-0.025	-0.025	1.491	0.000	0.146	0.200
1=no fun - 5=fun	3.6	4	1 - 5	3.8	4	1 - 5	-0.2	0	-3 - 1	-0.200	-0.200	1.005	0.000	0.329	0.000
1=long time - 5=quick	3.4	3.5	1 - 5	3.9	4	1 - 5	-0.5	-1	-4 - 2	-0.062	-0.938	1.318	-0.224	0.252	0.002

Table 5.3.2.12: The post-interview question difficulty questionnaire results of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Average question difficulty	1.795	1.8	1.3 - 2.7	1.66	1.6	1.3 - 2.8	0.135	0.2	-0.2 - 0.6	0.047	0.223	0.201	0.045	0.199	0.036
Do you have any pets?	1.15	1	1 - 2	1.1	1	1 - 2	0.05	0	-1 - 1	0.050	0.050	0.394	0.000	0.450	0.000
Why do people have pets?	2.45	2.5	1 - 3	2	2	1 - 3	0.45	0	0 - 1	0.450	0.450	0.510	0.000	0.361	0.000
What can children learn by having a pet?	2.05	2	1 - 3	2	2	1 - 3	0.05	0	-1 - 1	0.050	0.050	0.510	0.000	0.389	0.000
Can you tell me what a zoo is?	1.55	1	1 - 3	1.4	1	1 - 3	0.15	0	-1 - 1	0.150	0.150	0.489	0.000	0.420	0.000
What animals would you see at a zoo?	1.65	2	1 - 3	1.5	1.5	1 - 2	0.15	0	-1 - 1	0.150	0.150	0.587	0.000	0.351	0.000
What animal do you think is the most dangerous?	1.9	2	1 - 3	1.7	2	1 - 3	0.2	0	-1 - 2	0.200	0.200	1.056	0.000	0.225	0.009
What is the strangest animal you have seen?	2.2	2	1 - 3	1.75	2	1 - 3	0.45	0.5	-1 - 2	0.231	0.669	0.759	0.112	0.266	0.001
What is the largest animal you have seen?	1.3	1	1 - 3	1.65	1.5	1 - 3	-0.35	0	-2 - 1	-0.350	-0.350	0.671	0.000	0.349	0.000
Can you name some animals that live in this country?	2.25	2	1 - 3	2.15	2	1 - 3	0.1	0	-1 - 2	0.100	0.100	0.641	0.000	0.412	0.000
If you were an animal what animal would you like to be?	1.45	1	1 - 3	1.35	1	1 - 3	0.1	0	-1 - 1	0.100	0.100	0.553	0.000	0.372	0.000

Table 5.3.2.13: The question specific transcript interaction metrics of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Do you have any pets?	23.1	12	0 - 190	19.95	10	0 - 68	3.15	0	-68 - 127	3.150	3.150	36.638	0.000	0.226	0.009
Why do people have pets?	16.35	13	2 - 62	17.2	11	3 - 46	-0.85	-1	-25 - 21	-0.412	-1.288	13.076	-0.224	0.195	0.044
What can children learn by having a pet?	14.15	10.5	0 - 45	21.6	13.5	2 - 69	-7.45	-3.5	-33 - 13	-5.916	-8.984	13.748	-0.783	0.181	0.087
Can you tell me what a zoo is?	24.4	20.5	0 - 85	25	24	1 - 59	-0.6	-3	-26 - 56	0.715	-1.915	18.417	-0.671	0.237	0.004
What animals would you see at a zoo?	19.7	15	5 - 62	17.55	13	5 - 52	2.15	1	-43 - 42	1.712	2.588	16.863	0.224	0.210	0.021
What animal do you think is the most dangerous?	22.95	19	4 - 53	26.6	23.5	4 - 78	-3.65	-3.5	-43 - 35	-2.116	-5.184	16.490	-0.783	0.170	0.134
What is the strangest animal you have seen?	22.8	21	2 - 50	31.9	22	4 - 110	-9.1	-2	-78 - 20	-8.223	-9.977	24.458	-0.447	0.200	0.034
What is the largest animal you have seen?	14.15	4	1 - 55	9.4	2.5	1 - 47	4.75	0.5	-44 - 51	4.531	4.969	18.629	0.112	0.309	0.000
Can you name some animals that live in this country?	15.85	9	0 - 72	14.45	13	1 - 29	1.4	0	-20 - 63	1.400	1.400	18.048	0.000	0.207	0.024
If you were an animal what animal would you like to be?	32.4	25.5	5 - 142	24.95	22	3 - 78	7.45	3.5	-39 - 125	5.916	8.984	31.396	0.783	0.295	0.000

Table 5.3.2.14: The question specific temporal interaction metrics of phase 1 compared to phase 2.

Measure	Phase 1			Phase 2			Difference (Phase 1 – Phase 2)								
	Mean	Median	Range	Mean	Median	Range	Mean	Median	Range	Con. Lo	Con. Up	Standard Dev.	Standard Err.	K-S stat	P stat
Do you have any pets?	00:20	00:17	00:00 - 01:38	00:21	00:17	00:00 - 00:47	-00:01	-00:00	-00:45 - 00:58	-00:01	-00:01	00:24	-00:00	0.197	0.041
Why do people have pets?	00:20	00:17	00:08 - 00:45	00:23	00:20	00:06 - 00:48	-00:03	-00:02	-00:32 - 00:19	-00:02	-00:04	00:14	-00:01	0.148	0.200
What can children learn by having a pet?	00:17	00:16	00:10 - 00:32	00:21	00:19	00:05 - 00:43	-00:03	-00:03	-00:14 - 00:12	-00:02	-00:04	00:08	-00:01	0.102	0.200
Can you tell me what a zoo is?	00:19	00:16	00:07 - 00:45	00:20	00:21	00:07 - 00:38	-00:01	-00:02	-00:15 - 00:21	-00:00	-00:02	00:09	-00:01	0.114	0.200
What animals would you see at a zoo?	00:35	00:35	00:20 - 00:56	00:33	00:29	00:18 - 01:35	00:02	00:05	-00:48 - 00:29	-00:00	00:05	00:15	00:01	0.192	0.052
What animal do you think is the most dangerous?	00:31	00:32	00:14 - 00:54	00:29	00:27	00:12 - 00:49	00:02	00:01	-00:17 - 00:26	00:02	00:02	00:12	00:00	0.091	0.200
What is the strangest animal you have seen?	00:33	00:30	00:16 - 01:05	00:37	00:36	00:15 - 01:46	-00:04	-00:04	-01:02 - 00:50	-00:02	-00:06	00:21	-00:01	0.175	0.111
What is the largest animal you have seen?	00:21	00:15	00:06 - 00:51	00:17	00:12	00:07 - 00:44	00:04	00:03	-00:30 - 00:32	00:03	00:05	00:14	00:01	0.157	0.200
Can you name some animals that live in this country?	00:36	00:36	00:11 - 01:01	00:35	00:33	00:10 - 01:12	00:02	00:06	-00:38 - 00:40	-00:01	00:04	00:20	00:01	0.131	0.200
If you were an animal what animal would you like to be?	00:30	00:26	00:15 - 01:20	00:26	00:26	00:11 - 00:44	00:04	00:02	-00:22 - 01:01	00:03	00:05	00:18	00:00	0.149	0.200

5.3.3. Data visualisations

The boxplots presents in this section represent the differences between the robotic interviewer and the human interviewer in each of the measures stated.

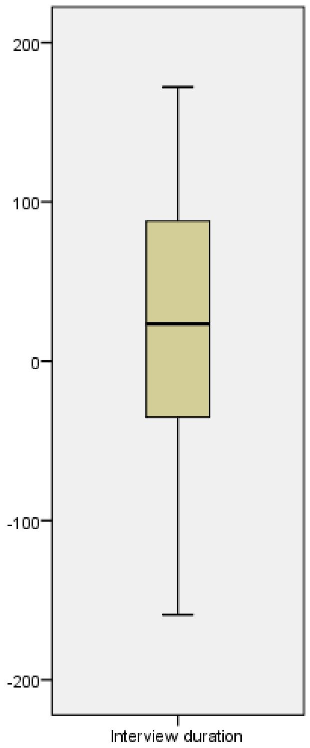


Figure 5.3.3.1: (seconds)

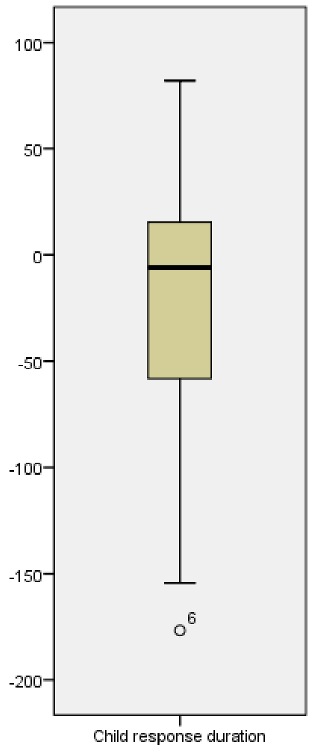


Figure 5.3.3.2: (seconds)

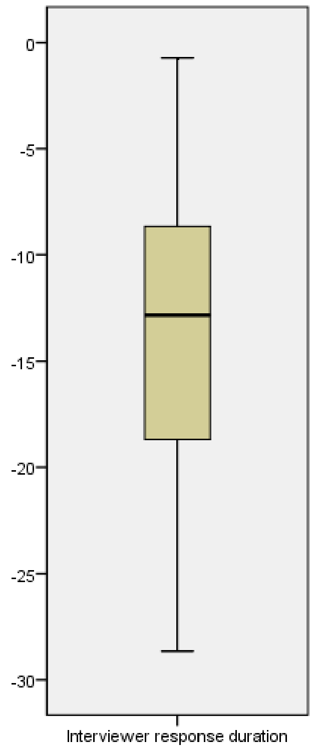


Figure 5.3.3.3: (seconds)

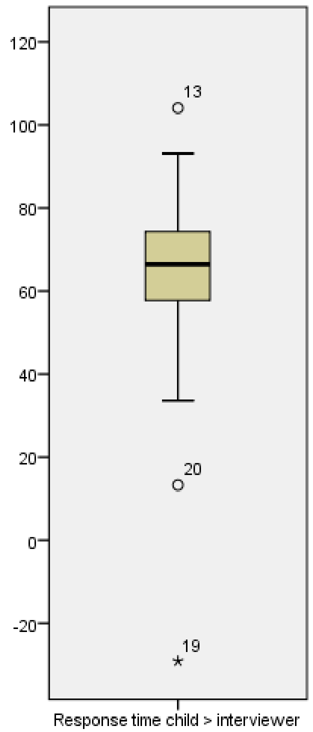


Figure 5.3.3.4: (seconds)

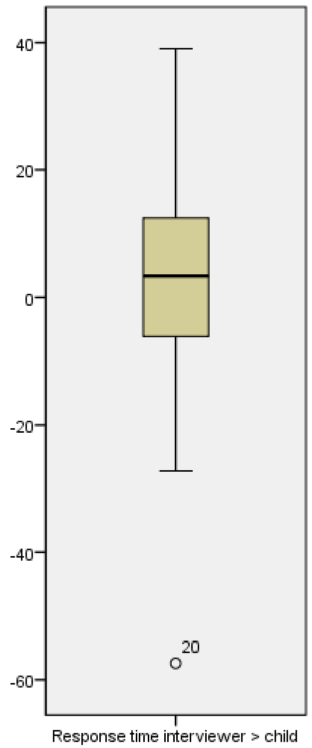


Figure 5.3.3.5: (seconds)

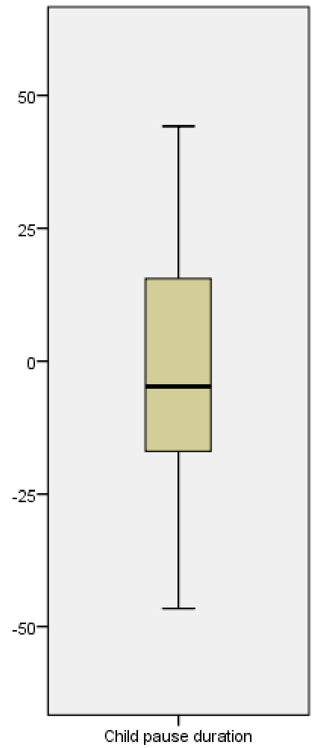


Figure 5.3.3.6: (seconds)

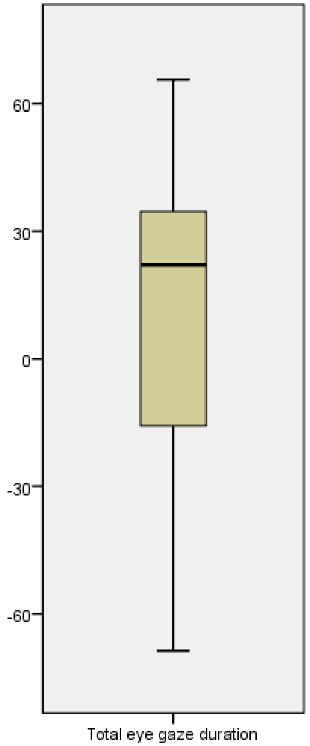


Figure 5.3.3.7: (seconds)

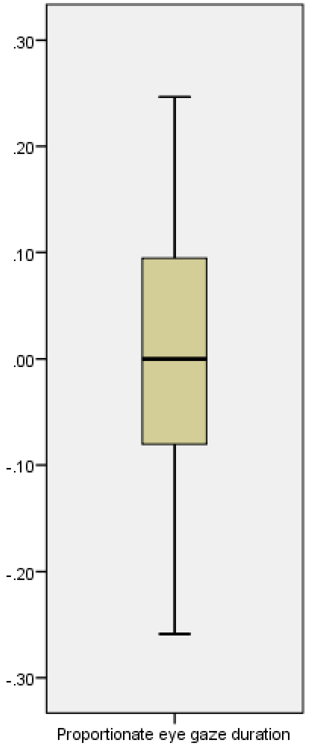


Figure 5.3.3.8: (eye gaze / duration)

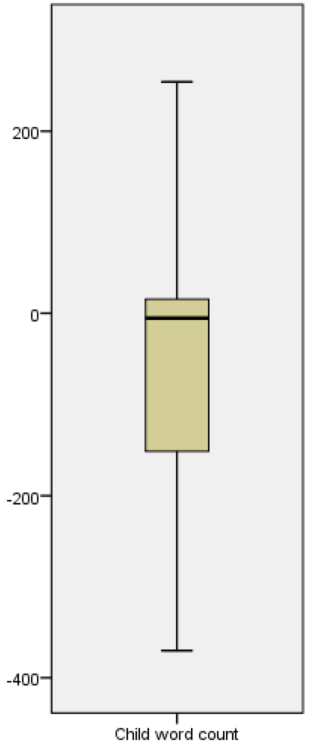


Figure 5.3.3.9: (words)

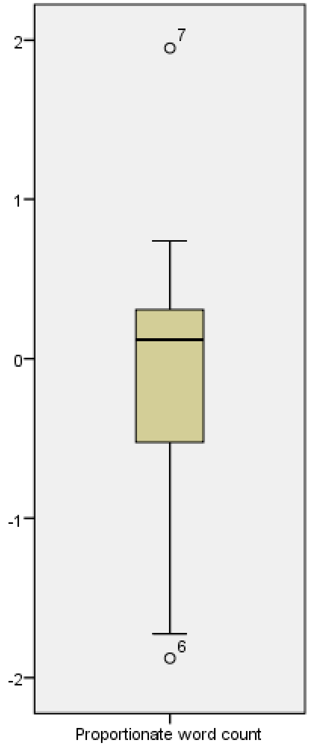


Figure 5.3.3.10: (child word count / interviewer word count)

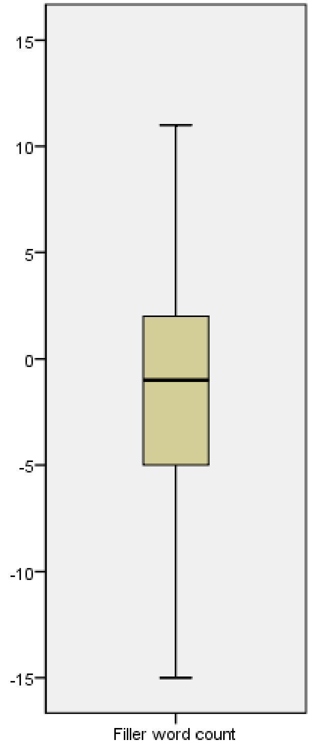


Figure 5.3.3.11: (words)

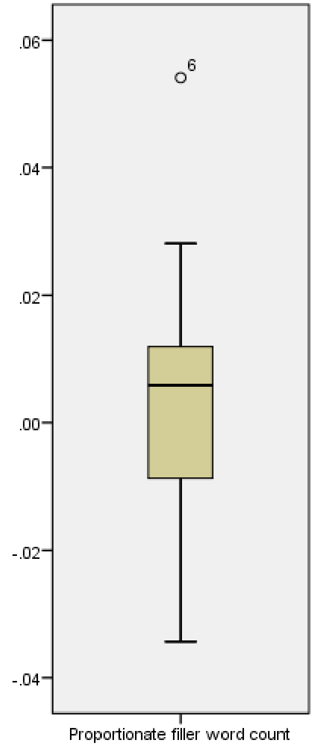


Figure 5.3.3.12: (filler word count / word count)

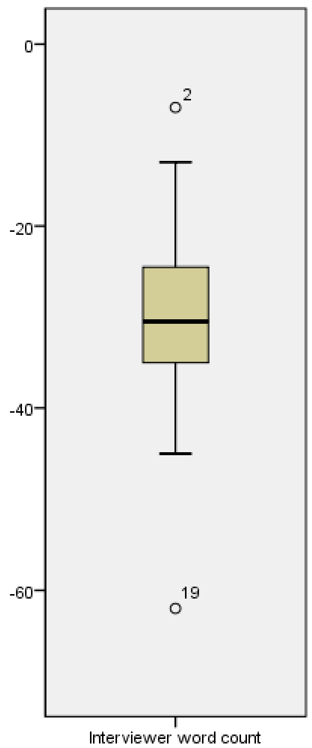


Figure 5.3.3.13: (words)

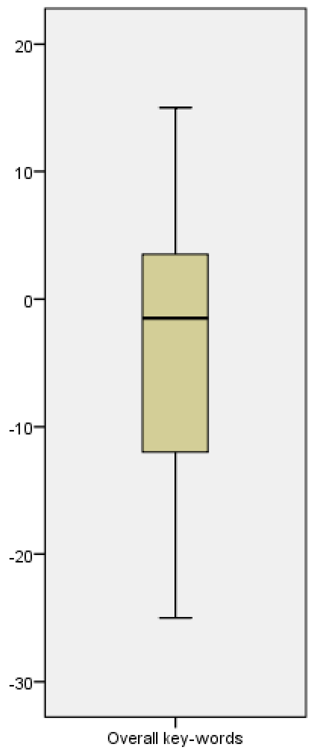


Figure 5.3.3.14: (key-words)

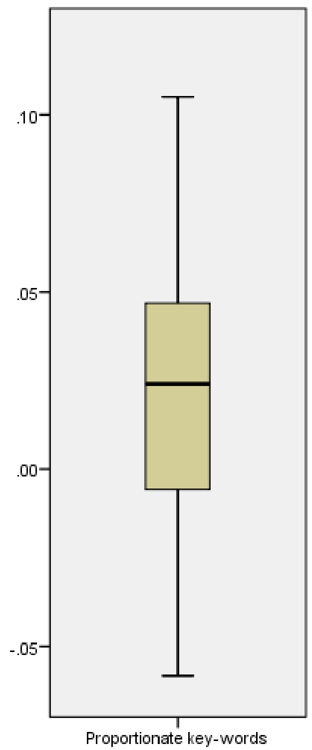


Figure 5.3.3.15: (key-words / word count)

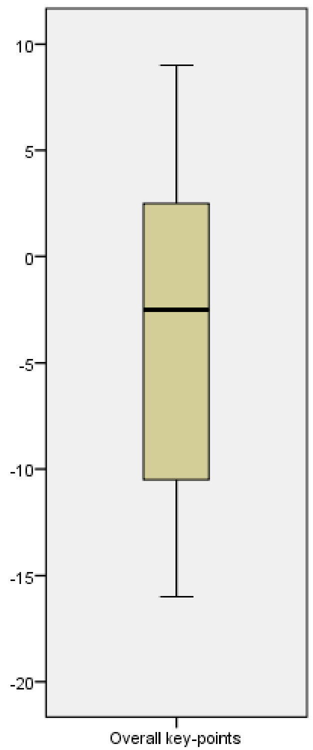


Figure 5.3.3.16: (key-points)

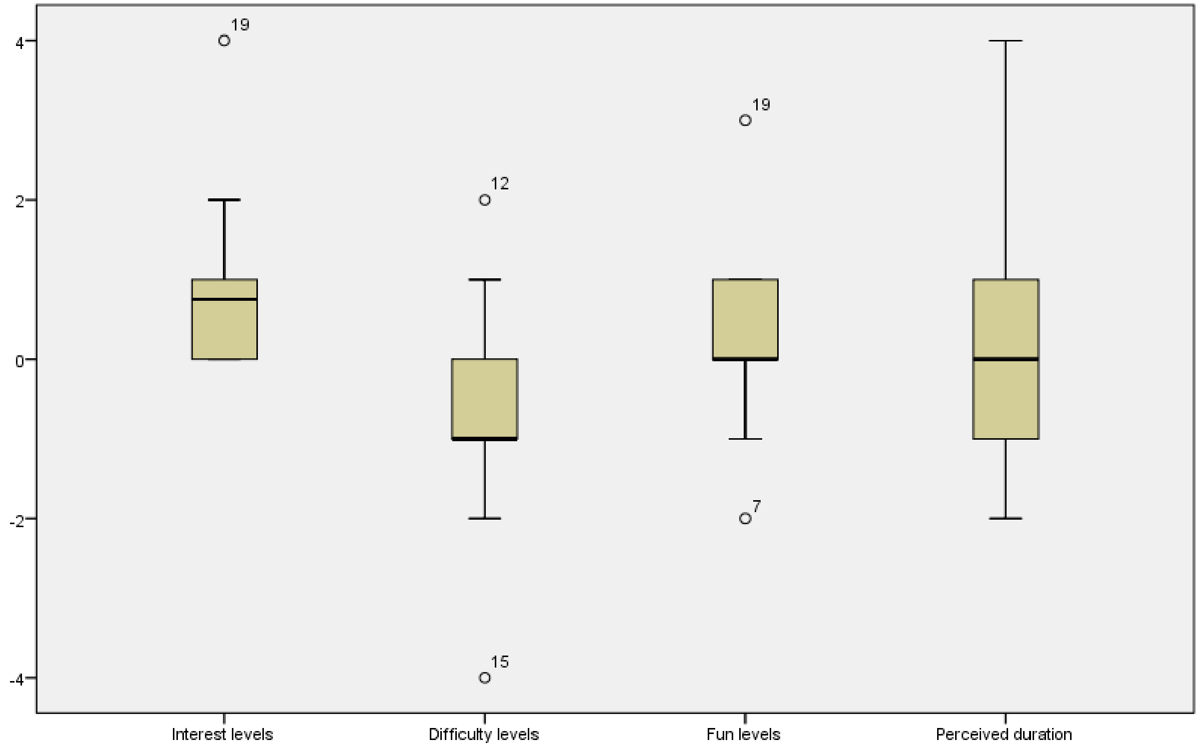


Figure 5.3.3.17: Questionnaire ratings

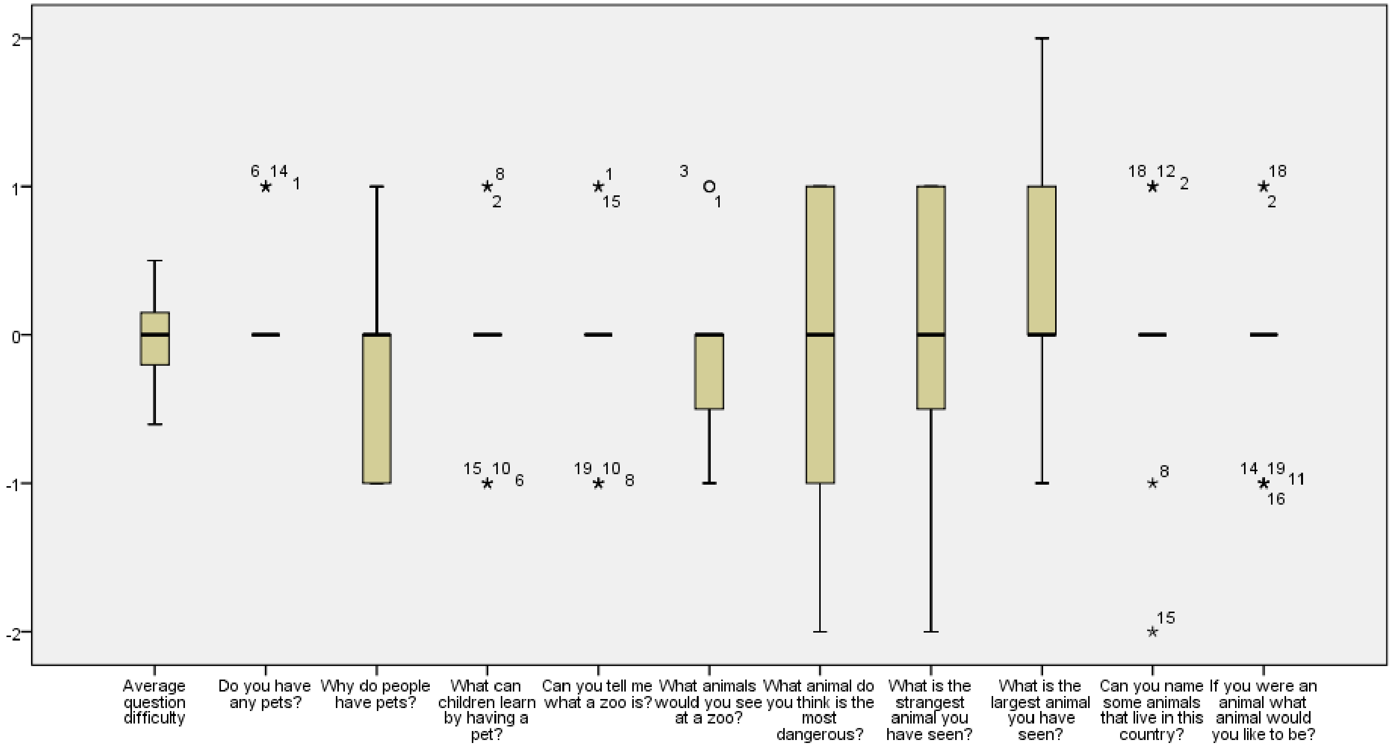


Figure 5.3.3.18: Question difficulty ratings

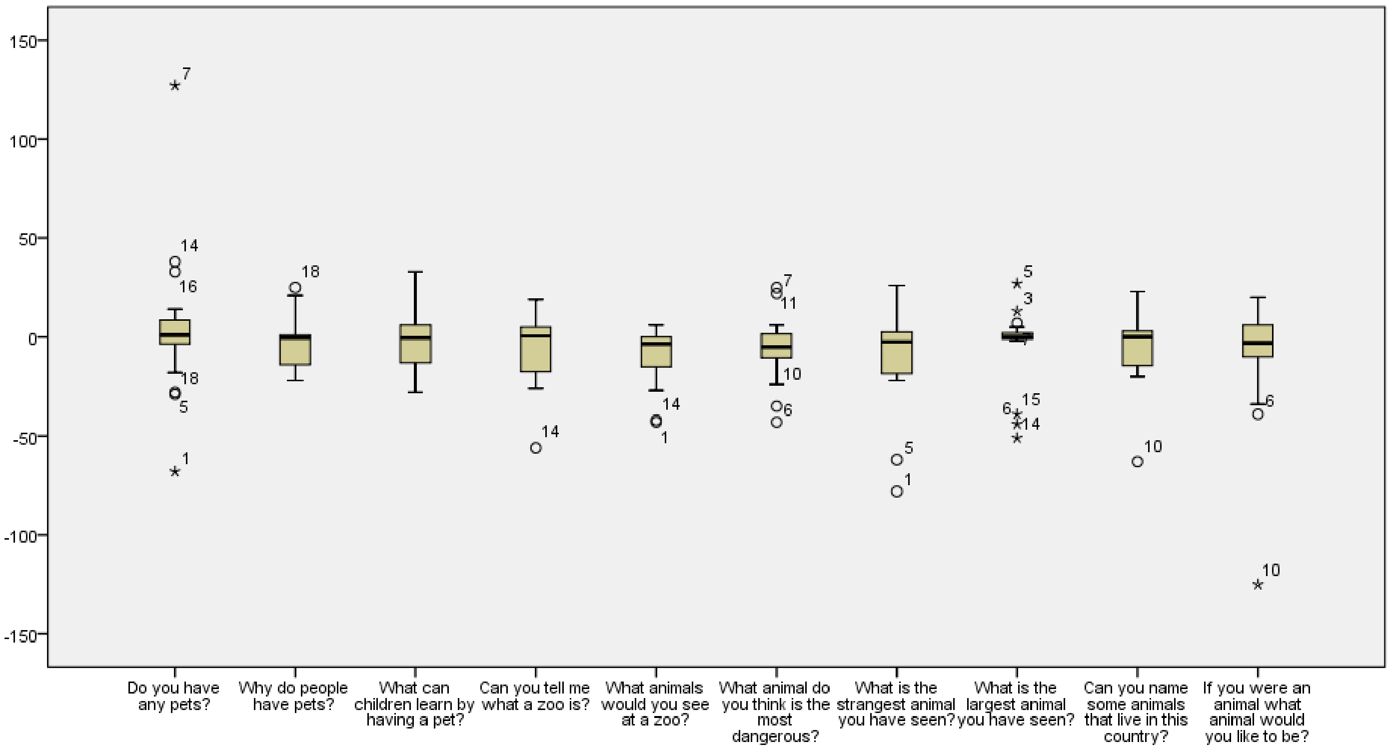


Figure 5.3.3.19: Question specific word counts

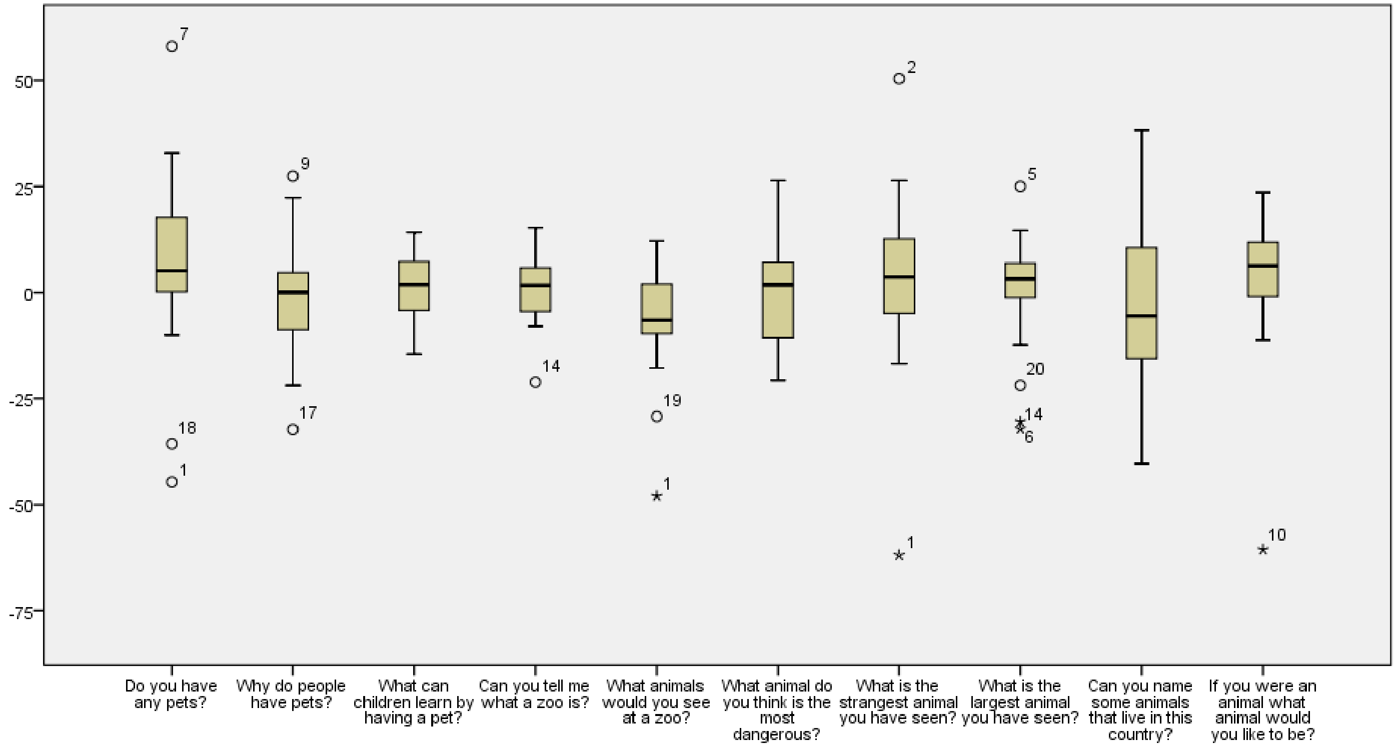


Figure 5.3.3.20: Question specific durations

5.3.4. Reliability of coded data

As stated in Section 3.3.1.4.3, 20% of the video data was coded by an independent reliability coder. The reliability coding in this study was conducted by Maria Jesus Hervas Garcia and a kappa of 0.86 was achieved, which is generally considered very good. This result indicates that the coding in this study was conducted to a high standard and the data produced from this coding is therefore reliable.

5.4. Discussion

5.4.1. Summary of findings

The findings confirm that the questions presented to the children varied in difficulty with an even distribution of easy, medium and hard questions. The results suggest KASPAR generally had very little impact on the behavior or information the children provide, see Tables 5.3.2.3 and 5.3.2.4. Furthermore, how difficult the children found the questions did not appear to vary between interviewers, see Table 5.3.2.6. Investigating the median word counts for the ten primary questions it was found that the amount that the children spoke to the human and robotic interviewer was very similar for the majority of the questions, see Tables 5.3.2.7. The most notable differences from the questionnaire data was how interesting the children found the activity and how difficult they found speaking to the interviewer. The median difference in Table 5.3.2.5 suggested that the children found the activity more fun with the robot, which is likely to be because talking to a robot is more novel than talking to a human. However, the children found talking to the robot more difficult, which was likely due to the robots text to speech synthesized voice. Table 5.3.2.4 shows that the children did reveal 2.5 key-points less to KASPAR than the human interviewer according to the median difference. Although this could suggest that the robotic interviewer may not be as effective as a human interviewer, the RMI in this particular instance was much more limited due to the technical limitations of the system, and the difference was very small.

5.4.2. Relation to literature and advances

This study has confirmed and reinforced the results of the previous study by showing that children do respond to a humanoid robot in an interview setting. This work also builds on Bethel et al.'s research by indicating that question difficulty does not appear to impact on the children's responses. In Bethel et al.'s study, the children were asked to keep a secret, and it was found that the children revealed this secret to the interviewer in every instance regardless of the interviewer [13]. However, in Bethel et al.'s study the focus was on eliciting one particular piece of information from the child, with no control for question difficulty. In this study a range of questions were asked that both varied in difficulty and were designed to elicit different types and amounts of information. In police investigations, it is important to secure as much information as possible because no single piece of information can be used on its own to build a case. This study along with Bethel et al.'s work begins to answer the questions that need addressing before RMI systems are trialed in a real world setting.

5.4.3. Limitations

Although this study has investigated how children respond to different types and difficulties of questions from a robot interviewer compared to a human interviewer, there are some limitations, particularly with regards to the type of questions. The questions that the children were being asked were not of a sensitive or personal nature which would likely have an effect on the children and the interactions if they were. Such type of questions were out of the scope of this thesis, and would require a different approach by specially trained experts in order to be able to obtain ethics approval[7]. For this scientific investigation, it is important to first conduct studies in less sensitive situations and develop the technology further before trying a RMI approach with such questions and settings.

5.4.4. Review of research questions and hypotheses

The results from this study support the previous study in Chapter 4, providing further data to answer the first two research questions: "RQ-1: To what degree do children respond to a robot in an interview setting?" and "RQ-2: How do children respond to a robot compared to a human interviewer?". In addition to this we have answered the third research question: "RQ-3: Will question type and difficulty impact on children's responses to a robot in an interview setting?". By finding that a robotic interviewer does not appear to impact perceived question difficulty or information disclosure regardless of question difficulty. The results of this study are in line with the initial hypothesis that the children will perform similarly with a robotic interviewer as they do a human interviewer.

5.4.5. Research questions for next study

The results from this study were consistent with the findings of the previous study and indicate that children were willing to interact with a robot in an interview setting and did so in a similar way to how they interacted with a human interviewer, providing additional data to answer the first two research questions, "RQ-1: To what degree do children respond to a robot in an interview setting?"

[7] On a personal note, the PhD candidate and his supervisors were of the opinion that it is morally questionable to conduct studies with child participants where they are exposed to personally sensitive or emotionally difficult questions.

and “RQ-2: How do children respond to a robot compared to a human interviewer?”. Further to this, the results of the study indicated that question difficulty does not appear to impact greatly on the amount of information that the children will provide to a robot, answering the third research question, “RQ-3: Will question type and difficulty impact on children’s responses to a robot in an interview setting?”. The results of this study again support the hypothesis that humanoid robots such as KASPAR could be useful tools for interviewing young children, and contributes to answer the overall primary research question, “Could robots be used as intermediaries for facilitating communication with children in an interview setting?”. Although the first two studies have investigated how neurotypical children respond to a robot, the literature review indicated that children with special needs are a group that may benefit from a RMI approach, possibly more than neurotypical children, therefore the next study will aim to answer the fourth research question: “RQ-4: How do children with special needs respond to a robot compared to a human interviewer?”. With children that have special needs being more likely to fall victim to maltreatment than neurotypical children [5], it would seem logical to investigate how children with special needs respond to a robot that is interviewing them, therefore this was the aim of the next study.

Chapter 6: How do children with special needs respond to a robotic interviewer (Study #3)

6.1. Introduction

The first two studies indicated that neurotypical children respond to a robotic interviewer in a very similar manner to which they do a human interviewer. However, research has shown that a particularly relevant target user group for RMI may be children with special needs. Jones et al. [5] conducted a systematic review of 17 papers and concluded that children with a disability are up to four times more likely to be a victim of abuse than children without disabilities, however the number of cases resulting in prosecution are relatively low [11, 12]. Children with special needs such as autism can sometimes be more difficult to communicate with, particularly when communicating with someone unfamiliar to them [96, 97]. These communication difficulties can sometimes be very obstructive when trying to acquire information from a child about sensitive or emotionally provocative events. A Mencap review states that "Those who can't communicate can't tell, and those who can't communicate well won't be believed" [6].

Also, research has shown that children with special needs such as autism are often very keen and willing to interact with robots [9, 98]. Robots such as KASPAR have been used for therapeutic purposes successfully on many occasions and have encouraged some of these children to be less isolated and more socially interactive [3, 4]. Using robots in this context for children with special needs is a logical step in light of the positive research indicating that children with special needs such as autism do respond well to robots in other contexts, see Chapter 2 for detailed background.

The study in this chapter investigated on a case study basis, how children with a variety of special needs respond to a robot in an interview setting compared to a human interviewer. We investigated how five children with special needs aged 9 to 11 responded to the KASPAR robot compared to a human in an interview setting. The five children were interviewed twice in a counterbalanced study similar to the first two studies. The standardised measures outlined in Section 3.3 were used[8] along with an additional measure of perceived engagement. Similar to the second study in Chapter 5, the main questions in the interviews varied in difficulty and focused on the theme of animals and pets. The results from quantitative data analysis were mainly consistent with the previous studies and indicted that the children interacted with KASPAR in a very similar manner to how they interacted

[8] Eye gaze was not measured in this study

with the human interviewer. Three of the children with specific communication needs appeared to be more engaged with the robotic interviewer and use more words whilst speaking to the robotic interviewer. However, this was counterbalanced by the majority of the children revealing less key-points to KASPAR than the human interviewer.

6.1.1. Research questions

The previous two studies in Chapters 4 and 5 have demonstrated that neurotypical children respond to a robotic interviewer in a very similar manner compared to how they respond to a human interviewer regardless of questions difficulty. These studies have contributed in answering the first three research questions: "RQ-1: To what degree do children respond to a robot in an interview setting?", "RQ-2: How do children respond to a robot compared to a human interviewer?" and "RQ-3: Will question type and difficulty impact on children's responses to a robot in an interview setting?". In this chapter we report on the third study investigating the possibility of RMIs with children that have various special needs in an effort to answer the fourth research question: "RQ-4: How do children with special needs respond to a robot compared to a human interviewer?".

6.1.2. Hypotheses

In light of the findings from the previous two studies and HRI literature, we would expect to see a clear preference towards the robotic interview partner, although this may not affect the amount of information that the children provide in this particular study. We predicted that the children would have a particularly clear preference to the robot with regards to how engaged they are with the interview partner and activity.

6.2. Method

6.2.1. Participants

This study took place in a UK primary school that specialises in catering for children with special needs. Five children (3 male, 2 female) took part in this study, each of which had different special needs that included: ASC (Autism Spectrum Condition), BESD (Behavioural, Emotional and Social Difficulties), MLD (Moderate Learning Difficulties), and SLCN (Speech, Language and Communication Needs), Table 6.2.1.1 shows which conditions each child had. The children were between the ages of 9 and 11. Some of the children had interacted with KASPAR in previous studies, however these studies were different in their nature and context. In accordance with the ethics procedure, informed consent was obtained in writing from the parents of all the children participating in the study as outlined in Section 3.1.2.3. The ethics protocol number for this study is 1213/17, the relevant documentation can be seen in Appendix 4.1.

Table 6.2.1.1: Participant condition table

Participant ID	Conditions	Codes – Conditions
1	ASC, BESD, MLD, SLCN	**ASC** - Autism Spectrum Condition **BESD** - Behavioural, Emotional and Social Difficulties **MLD** - Moderate Learning Difficulties **SLCN** - Speech, Language and Communication Needs
2	BESD, MLD	
3	ASC, MLD, SLCN	
4	MLD, SLCN	
5	MLD	

6.2.2. Procedure

As with the previous two studies the children received a group introduction to both KASPAR (Figure 3.2.2.1) and the human interviewer at the school one day prior to the interviews commencing. The group introduction was used to familiarise the children with the interviewers and explain the procedure to them. Similar to the previous studies each child had two interviews, one with the robot and one with the human interviewer. These interviews were carried out in a two phase counterbalanced structure as described in Section 3.2.7 to minimise any crossover effects. Due to time constraints the gap between each of the interviews was one day rather than a week. The interviews in this study followed the same structure as the previous studies and was based on the recommendations of the ABE [7]. The questions for this study were the same as the second study and were focused on the theme of animals and pets. This topic was chosen because it offered sufficient scope for various different types and difficulty of questions, and had been tested successfully in the previous study. Prior to the study commencing the deputy head teacher of the school was shown the proposed interview questions to establish if the questions were appropriate and select the children that would be most appropriate for the study. Because the school caters for children with a range of special needs, not all of the children in the school have verbal communication, therefore it was necessary to first establish which children would be appropriate for this study. After each interview the child was asked some questions about the experience, in particular the interviewer. Once all of the interviews had been completed, the children were given the opportunity to play with KASPAR along with a selection of other robotic toys.

6.2.3. The setup

The interviews were conducted in an unused classroom that contained a small lockable cupboard which the children were unable to see into. The cupboard was used as a control room for KASPAR and housed a monitor with a wireless connection to camera #1 to observe the situation and make

KASPAR respond appropriately, see Figure 6.2.3.1. The children were unaware that KASPAR was being controlled by a human triggering the correct questions and responses from a pre-recorded list. Again the interviews were led by the principle investigator either in person or remotely via KASPAR to help maintain consistency between the interviews. The children were taken to and from the interviews by a second researcher that was unknown to the children. The second researcher remained in the room throughout the interviews, but was as non-reactive as possible. After each interview the children were asked by the second researcher to rate particular aspects of the experience, in particular the interaction partner.

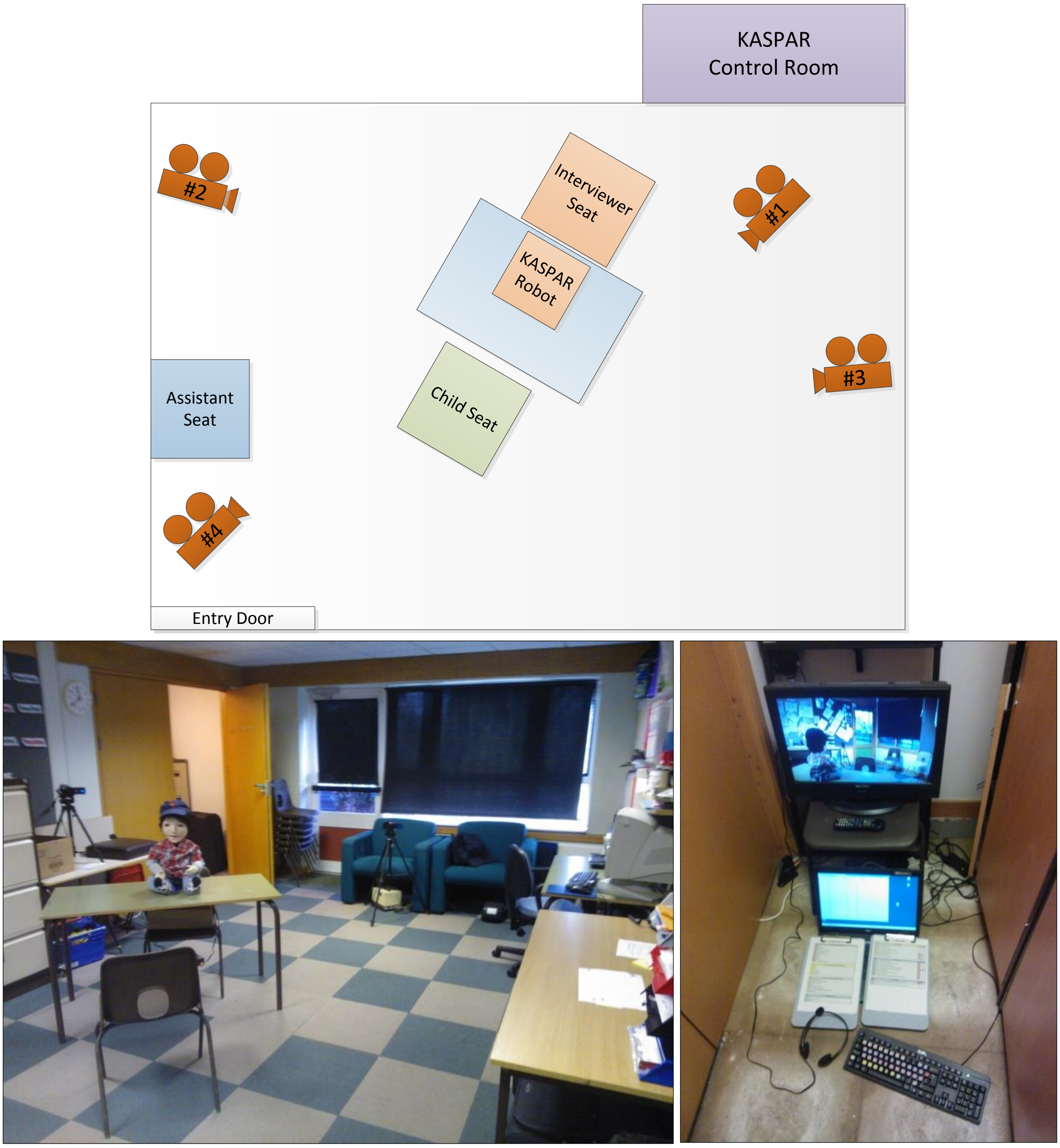


Figure 6.2.3.1: An overview of the room layout and interview setup

6.2.4. Measures

There were four primary units of analysis in this study, three of these were the standard techniques outlined in Section 3.3, however there was a fourth unit of analysis that we implemented for this study which was perceived engagement.

6.2.4.1. Perceived Engagement

In this study we analysed how engaged the children appeared to be in the activity and interaction partner. The engagement was coded by 3 independent blind coders that had no prior knowledge of the study or this area of research. The coders worked individually and were unknown to each other. Each coder viewed all of the videos from start to finish before commencing the coding in order to acquire some familiarity for the children and their behaviour first. This was an important phase because accurately assessing a child's engagement relies on having some previous knowledge of the child's full repertoire of behaviour within that particular context. The videos had been edited to remove the audio and obscure a portion of the screen to hide the interaction partner from the coders. The coders were advised of the interaction partner's location to assist in the coding process, but were not provided with any details of the experiment, as this may have adversely altered and biased their coding. In order to code the engagement a custom program was designed and built in Adobe Flash to display the video and record the coders input. The videos were coded with a simple on screen slider that moved with the position of the curser. A smile-meter placed in the top right hand corner of the screen changed according to the position of the slider. If the slider was towards the top this would indicate more engaged and the smile-meter would have a greater smile, if the slider was lower down the screen this would indicate less engaged and the smile-meter would change to more of a frown, see Figure 6.2.4.1.1. This approach of continuous monitoring was originally employed for market research [99], but has also been adopted and adapted for measuring levels of fun and engagement for HRI and HCI applications [17, 100]. The method used in this study takes inspiration from these approaches, but uses a slightly different implementation to suit the requirements of this particular type of study.

Figure 6.2.4.1.1: Screenshot of engagement coding program

The coders were given brief instructions on what may constitute engagement, see Appendix 4.2 for briefing sheet provided to coders. However, it was ultimately left to the coder to decide what they personally considered to be engagement. The coders would move the slider up and down on a touch screen according to how engaged they personally felt that the child was in the interaction partner. The coding occurred at the normal video playing speed and the program logged where the slider was on the screen in relation to the position of the video. The program logged the position of the slider 10 times every second. The coding of each video took place in one continuous run with no pausing or rewinding.

Neurotypical people are often very good at assessing how engaged someone is in an activity or an interaction partner. Rather than trying to code specific behaviours in great detail such as eye contact or body gestures, we tried this more subjective method and combined the results of each coder to establish an overall average for the engagement in each interview.

6.3. Results

In contrast to the previous 2 studies, we did not perform and overall analysis of the data, rather we look at the data on a case by case basis. Because there were only 5 children participating in this study and they all had different conditions, it was advantageous to look at the data on a case by case basis and present the results in Tables 6.3.2.1 to 6.3.2.6.

6.3.1. Data analysis

6.3.1.1. Temporal analysis

Examining the interview durations between the human and robotic interviewers we found that in 4 of the 5 interviews, when the children were speaking to KASPAR the interviews took more than 30 seconds longer, see Table 6.3.2.1. Investigating this more closely, although KASPAR takes much longer to respond, we can see that 3 of the children speak to KASPAR longer overall, see Table 6.3.2.1. Assessing these results against the children's conditions we can see that all three of these children have SLCN (Speech, Language and Communication Needs) and two of the children had autism, see Figure 6.2.1.1. This may provide some indication that an RMI approach may be more suitable with children that have SLCN and autism. Taking this hypothesis and looking at the temporal data further we can see another pattern. The children with SLCN and autism all spend more time pausing whilst speaking to KASPAR whilst the other 2 children pause less time whilst speaking to KASPAR.

6.3.1.2. Textual content analysis

Looking at the amount of words that the children spoke to each interviewer there appears to be a considerable difference in the number of words the children use between the interviewers depending on their conditions, see Table 6.3.2.2. The children with SLCN and autism appear to speak to KASPAR much more than the human interviewer, they also use more filler words and their proportionate word count to the interviews word count is also higher. However, the additional filler words and relative word counts are to be expected if the children use more words when speaking to KASPAR. This pattern continues with the number of key-words that the children use, with the children that have SLCN and autism using more key-words with KASPAR than the human interviewer. However, when analysing the key-words relative to the children's overall word count it appears that proportionally these children use less with KASPAR however, given that the children used more words when speaking to KASPAR this is logical. The analysis of the key-points reveals that 4 out of the 5 children revealed less key-points to KASPAR than the human interviewer, however this may have been because the children felt less pressure to provide information to the robot that the human interviewer. It would appear from the analysis of the textual content that the 3 children with SLCN and (2 of which also have autism) may have had a preferred the robotic interviewer, whilst the other children seemed to prefer the human interviewer. This shows that a RMI approach would likely be more useful for particular groups of children, and that this would need to be assessed.

6.3.1.3. Basic questionnaire analysis

The experience questionnaire results were mostly similar, with the most notable difference being how long the children felt that the interviews had taken. 4 of the 5 children felt that the interviews with KASPAR took longer, other than this the responses were mixed with a tendency to rate the questionnaires identically for both interviewers, see Table 6.3.2.3.

6.3.1.4. Question specific data

Taking a detailed look at the word counts and the time spent on the main interview questions, it can be seen that the children with SLCN and autism again seem to spend more time responding to KASPAR and use more words for the majority of the questions, whereas the other 2 children spend more time talking to the human interviewer, see Table 6.3.2.5. However, there was one question for which all of the children seemed to use less words when talking to KASPAR which was "Can you name some animals that live in this country?", Table 6.3.2.4. This is partially consistent with the previous study because the question required the children to list animals, this therefore further contributes to the theory that children may not respond as well to robots when presented with questions that require them to list.

6.3.1.5. Perceived engagement

The results from this study seem to indicate that the children with SLCN and autism may respond better to a robot than a human interviewer. Further to this, the qualitative results from looking at the interactions in the two conditions would indicate that some of the children seemed to be more engaged with the robot than the human interviewer, see Table 6.3.2.6 and Figure 6.3.3.1. From the five children that took part in the study it appears that the children with the most severe communication difficulties appeared to be a lot more engaged with the robot than the human interviewer. For example, three of the children's engagement scores were on average 108.99, 47.20 and 44.65 higher with KASPAR than the human interviewer. All three of these children had SLCN and two of the children had autism, which is consistent with the other results in this study. However, this data needs to be treated with caution. As can be seen from the graphs in Figures 6.3.3.3 to 6.3.3.7 although there are some trends and agreement in the data, there is a lot of inconstancies and variation, therefore this data can only be used to supplement the other measures as it would not be robust enough to hold value on its own.

6.3.1.5.1. Working with engagement data

The data gathered by the engagement methodology was noisy as can be seen in the raw data sets in Figures 6.3.3.3 to 6.3.3.7. In an effort to correct for errors arising from idiosyncrasies in regard to the timing and baseline engagement-levels for each coder, each coder's responses were standardised around their overall mean rating, and the temporal horizon was extended outwards by 30 seconds through computing a weighted moving average (exponential smoothing). This allows for an easier visual assessment of overall trends between the coders, see Figures 6.3.3.3 to 6.3.3.7. The results of the principle investigators coding were also included in the dataset for the reliability analysis but were not used for the main analysis. When assessing the agreement between the coders the smoothed and standardised data was used when computing the overall Intra-Class Correlation Coefficient (ICC) [101]. We found that there was a moderate but significant effect according to the Correlation Coefficients in Table 6.3.2.7. Future work could investigate this methodology in more detail.

6.3.2. Data tables

- Table 6.3.2.1 lists the individual temporal measurements for both the robotic and human interviewer with each child, and highlights the differences
- Table 6.3.2.2 lists the individual transcription measurements for both the robotic and human interviewer with each child, and highlights the differences
- Table 6.3.2.3 lists the individual basic post-interview questionnaire measurements for both the robotic and human interviewer with each child, and highlights the differences
- Table 6.3.2.4 lists the individual question specific transcription measurements for both the robotic and human interviewer with each child, and highlights the differences
- Table 6.3.2.5 lists the question specific temporal measurements for both the robotic and human interviewer with each child, and highlights the differences
- Table 6.3.2.6 lists the engagement scores for the children interacting with each interviewer and the differences between the interviewers
- Table 6.3.2.7 lists the results of the Correlation Coefficients for coded engagement data

Table 6.3.2.1: The temporal interaction metrics of KASPAR compared to the human interviewer.

Participant ID: (Condition codes)	1: (ASC, BESD, MLD, SLCN)			2: (BESD, MLD)			3: (ASC, MLD, SLCN)			4: (MLD, SLCN)			5: (MLD)		
	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference
Interview duration	06:26	04:59	01:28	04:22	03:52	00:30	06:30	03:19	03:11	04:41	03:40	01:01	05:24	05:55	-00:31
Child response duration	01:43	00:57	00:46	00:41	01:15	-00:34	03:38	01:17	02:21	02:02	00:51	01:11	01:00	02:15	-01:15
Interviewer response duration	01:00	01:23	-00:24	01:02	01:09	-00:07	00:47	01:09	-00:21	00:53	01:14	-00:21	00:54	01:29	-00:35
Response time child > interviewer	02:14	00:26	01:48	01:43	00:27	01:17	01:02	00:19	00:43	00:46	00:54	-00:09	02:04	00:48	01:16
Response time interviewer > child	01:09	00:43	00:25	00:36	00:47	-00:11	00:25	00:16	00:09	00:09	00:17	-00:08	00:44	00:28	00:16
Child pause duration	00:11	00:05	00:06	00:11	00:10	00:01	00:35	00:05	00:30	00:51	00:10	00:41	00:41	00:43	-00:02

Table 6.3.2.2: The transcript interaction metrics of KASPAR compared to the human interviewer.

Participant ID: (Condition codes)	1: (ASC, BESD, MLD, SLCN)			2: (BESD, MLD)			3: (ASC, MLD, SLCN)			4: (MLD, SLCN)			5: (MLD)		
	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference
Child word count	122	68	54	79	133	-54	415	152	263	98	51	47	65	112	-47
Proportionate word count	0.718	0.402	0.315	0.467	0.679	-0.211	3.051	0.927	2.125	0.616	0.280	0.336	0.422	0.580	-0.158
Filler word count	2	0	2	5	5	0	26	6	20	12	1	11	7	18	-11
Proportionate filler word count	0.016	0.000	0.016	0.063	0.038	0.026	0.063	0.039	0.023	0.122	0.020	0.103	0.108	0.161	-0.053
Interviewer word count	170	169	1	169	196	-27	136	164	-28	159	182	-23	154	193	-39
Overall key-words	14	13	1	12	19	-7	49	20	29	21	20	1	16	20	-4
Proportionate key-word count	0.115	0.191	-0.076	0.152	0.143	0.009	0.118	0.132	-0.014	0.214	0.392	-0.178	0.246	0.179	0.068
Overall key-points	10	14	-4	10	16	-6	37	21	16	23	25	-2	12	18	-6

Table 6.3.2.3: The post-interview questionnaire results of the interactions with KASPAR compared to the human interviewer.

Participant ID: (Condition codes)	1: (ASC, BESD, MLD, SLCN)			2: (BESD, MLD)			3: (ASC, MLD, SLCN)			4: (MLD, SLCN)			5: (MLD)		
	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference
1=boring - 5=interesting	5	5	0	3	3	0	4	3	1	1	5	-4	4	1	3
1=hard - 5=easy	1	1	0	5	5	0	3	4	-1	5	5	0	2	5	-3
1=no fun - 5=fun	5	5	0	5	5	0	4	3	1	5	5	0	4	1	3
1=long time - 5=quick	1	5	-4	5	5	0	3	5	-2	1	5	-4	1	5	-4

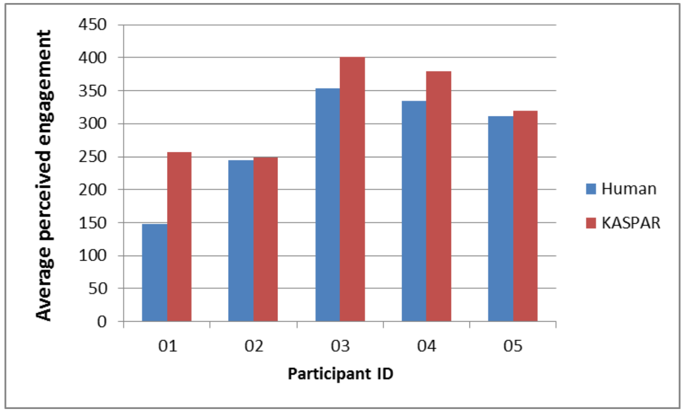
Table 6.3.2.4: The question specific transcript interaction metrics of KASPAR compared to the human interviewer.															
Participant ID: (Condition codes)	1: (ASC, BESD, MLD, SLCN)			2: (BESD, MLD)			3: (ASC, MLD, SLCN)			4: (MLD, SLCN)			5: (MLD)		
	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference
Do you have any pets?	10	5	5	1	1	0	92	15	77	1	7	-6	1	1	0
Why do people have pets?	18	11	7	3	6	-3	36	27	9	6	3	3	3	1	2
What can children learn by having a pet?	8	0	8	3	15	-12	41	5	36	5	2	3	5	2	3
Can you tell me what a zoo is?	2	13	-11	7	25	-18	38	7	31	7	1	6	1	6	-5
What animals would you see at a zoo?	5	2	3	2	8	-6	13	3	10	4	7	-3	1	9	-8
What animal do you think is the most dangerous?	31	7	24	12	6	6	27	15	12	7	3	4	8	5	3
What is the strangest animal you have seen?	10	10	0	3	3	0	14	8	6	9	2	7	3	5	-2
What is the largest animal you have seen?	9	2	7	5	14	-9	29	5	24	7	3	4	1	20	-19
Can you name some animals that live in this country?	3	5	-2	1	8	-7	15	20	-5	6	9	-3	2	7	-5
If you were an animal what animal would you like to be?	9	6	3	15	6	9	26	5	21	9	3	6	11	40	-29

Table 6.3.2.5: The question specific temporal interaction metrics of KASPAR compared to the human interviewer.															
Participant ID: (Condition codes)	1: (ASC, BESD, MLD, SLCN)			2: (BESD, MLD)			3: (ASC, MLD, SLCN)			4: (MLD, SLCN)			5: (MLD)		
	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference	KASPAR	Human	Difference
Do you have any pets?	00:37	00:23	00:14	00:05	00:06	-00:01	00:49	00:21	00:28	00:05	00:22	-00:17	00:06	00:05	00:01
Why do people have pets?	00:30	00:23	00:07	00:09	00:13	-00:04	00:23	00:24	-00:01	00:16	00:06	00:10	00:17	00:21	-00:05
What can children learn by having a pet?	00:19	00:28	-00:09	00:11	00:23	-00:12	00:33	00:11	00:22	00:16	00:17	-00:01	00:25	00:21	00:04
Can you tell me what a zoo is?	00:26	00:18	00:08	00:11	00:21	-00:10	00:27	00:09	00:17	00:17	00:14	00:03	00:13	00:18	-00:05
What animals would you see at a zoo?	00:27	00:15	00:12	00:27	00:17	00:11	00:28	00:07	00:22	00:19	00:21	-00:03	00:28	00:24	00:03
What animal do you think is the most dangerous?	00:37	00:18	00:19	00:29	00:10	00:18	00:29	00:17	00:12	00:24	00:11	00:13	00:32	00:40	-00:07
What is the strangest animal you have seen?	00:31	00:27	00:03	00:11	00:08	00:03	00:21	00:10	00:11	00:17	00:11	00:06	00:27	00:26	00:00
What is the largest animal you have seen?	00:19	00:12	00:06	00:22	00:24	-00:02	00:31	00:18	00:13	00:20	00:17	00:03	00:17	00:58	-00:41
Can you name some animals that live in this country?	00:23	00:26	-00:03	00:24	00:21	00:04	00:33	00:25	00:08	00:13	00:27	-00:14	00:11	00:29	-00:17
If you were an animal what animal would you like to be?	00:55	00:21	00:34	00:30	00:15	00:15	00:36	00:09	00:27	00:23	00:14	00:09	00:43	00:48	-00:05

Table 6.3.2.6: Average perceived engagement scores of the children interacting with each interviewer				
Participant ID	**Conditions**	**Human**	**KASPAR**	**Difference**
1	ASC, BESD, MLD, SLCN	147.316	256.311	108.994
2	BESD, MLD	244.319	248.703	4.384
3	ASC, MLD, SLCN	354.039	401.242	47.203
4	MLD, SLCN	335.115	379.763	44.649
5	MLD	310.655	319.453	8.798

Table 6.3.2.7: Correlation Coefficients for engagement data		
	ICC	**95% CI**
Raw Ratings	0.05	0.04-0.05
Standardised Ratings	-0.13	-0.13--0.12
Smoothed Standardised Ratings	-0.24	-0.25--0.24

6.3.3. Data visualisations



Codes – Conditions

ASC - Autism Spectrum Condition

BESD - Behavioural, Emotional and Social Difficulties

MLD - Moderate Learning Difficulties

SLCN - Speech, Language and Communication Needs

Figure 6.3.3.1: Average perceived engagement graph

Observer	
	Lead investigator
	Independent coder 1
	Independent coder 2
	Independent coder 3

Figure 6.3.3.2: Observer legend

Figure 6.3.3.3: Participant 1 Engaement data

Figure 6.3.3.4: Participant 2 Engaement data

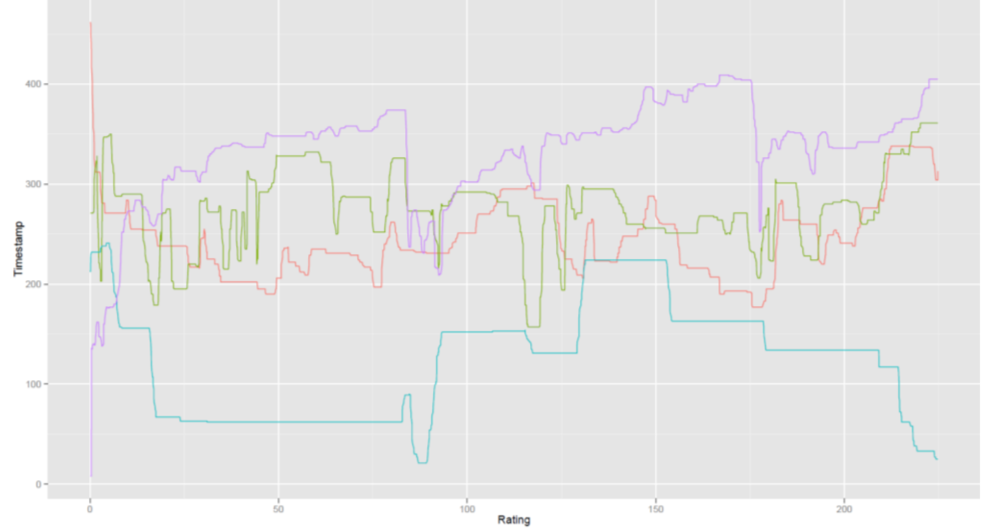
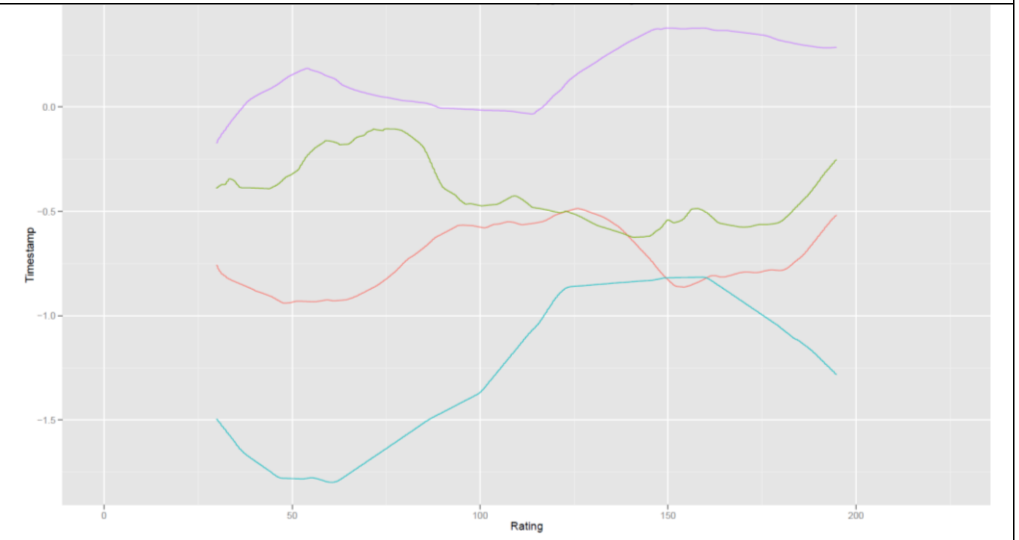
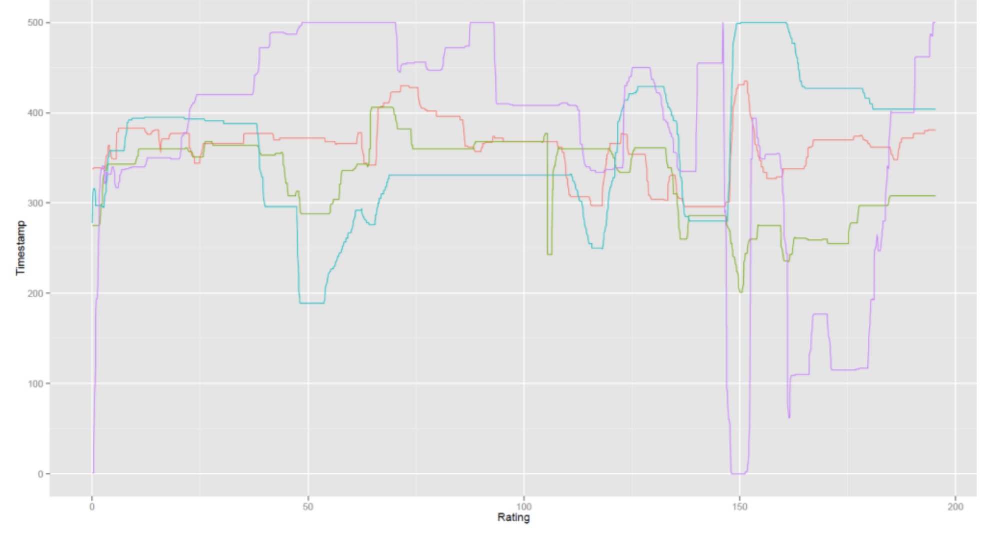
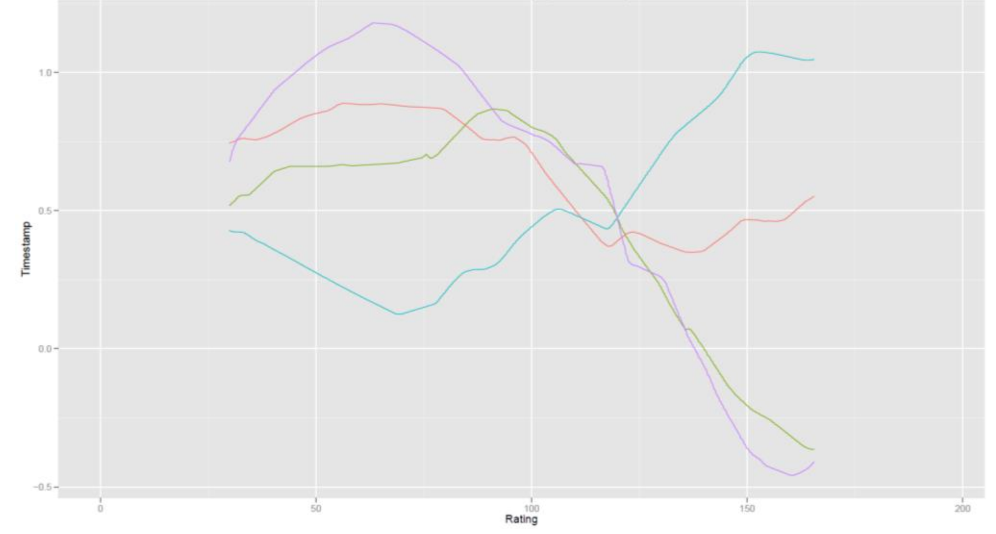
	Raw Engagement Ratings	Smoothed Standardised Engagement Ratings
KASPAR		
Human		

Figure 6.3.3.5: Participant 3 Engaement data

Figure 6.3.3.6: Participant 4 Engaement data

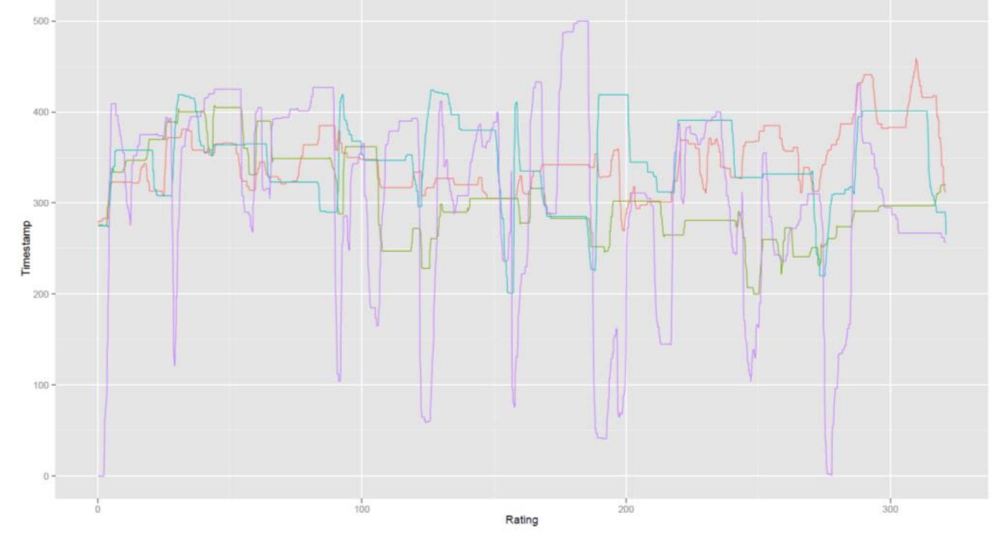
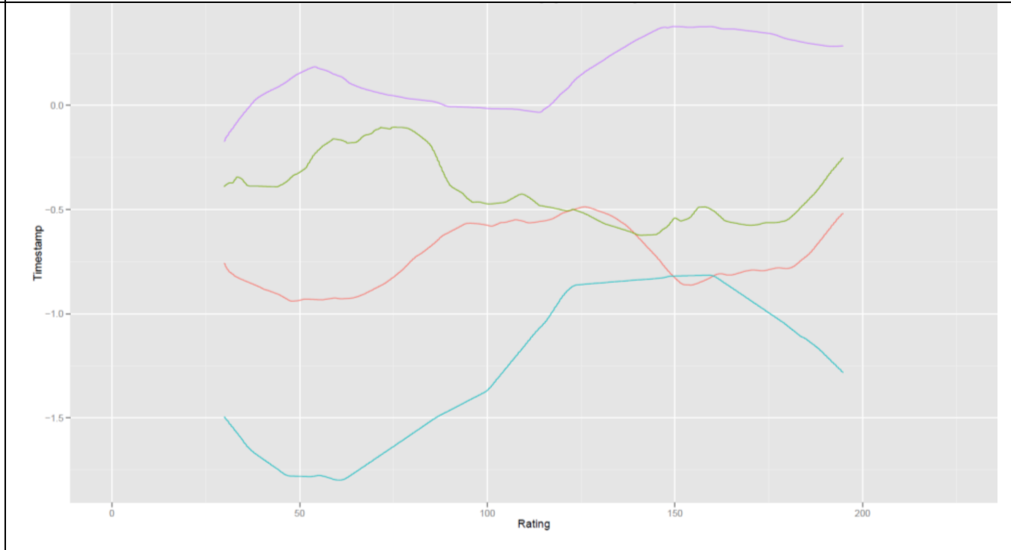
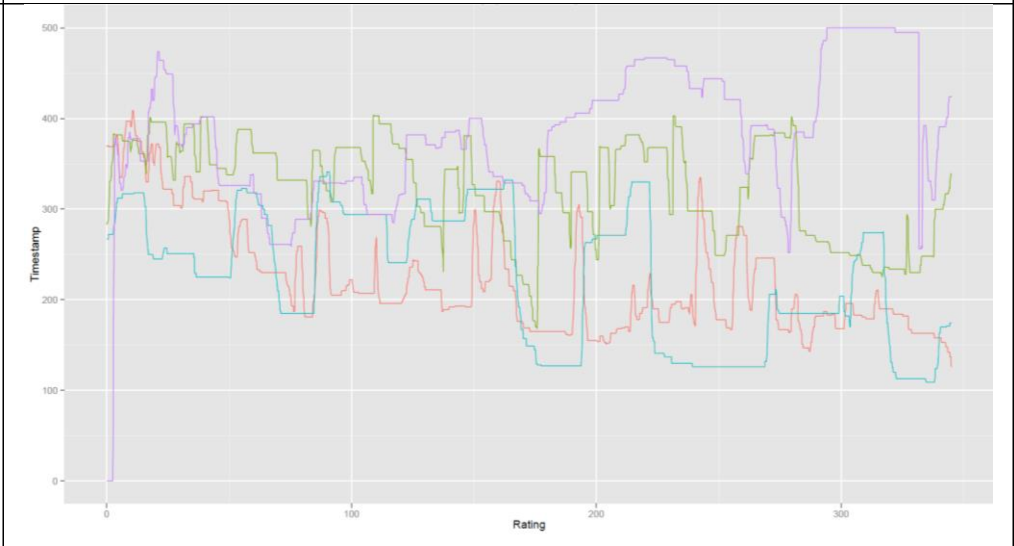
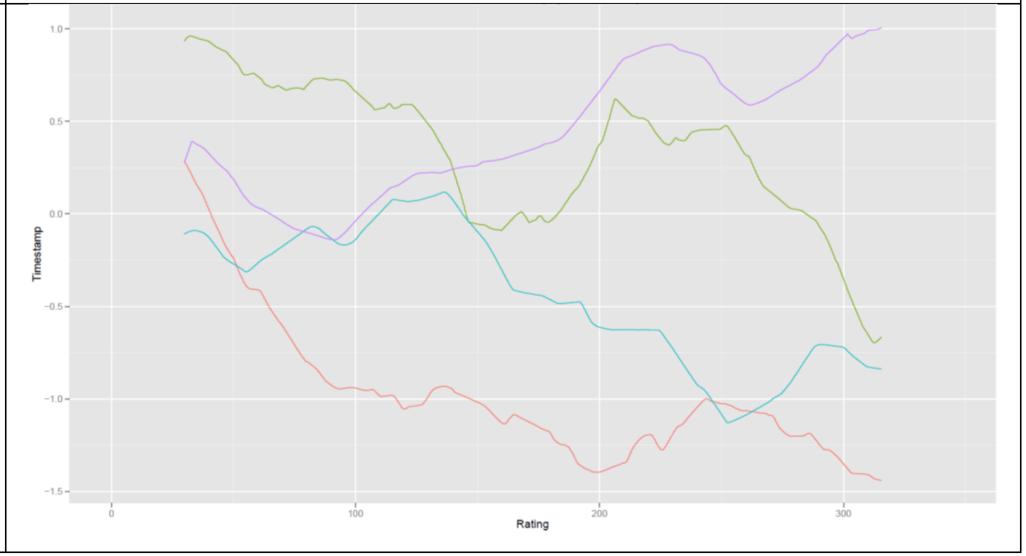
	Raw Engagement Ratings	Smoothed Standardised Engagement Ratings
KASPAR		
Human		

Figure 6.3.3.7: Participant 5 Engaement data

6.3.4. Reliability of coded data

As stated in Section 3.3.1.4.3, 20% of the video data was coded by an independent reliability coder. The reliability coding in this study was conducted by Maria Jesus Hervas Garcia and a kappa of 0.80 was achieved, which is generally considered very good. This result indicates that the coding in this study was conducted to a high standard and the data produced from this coding is therefore reliable.

6.4. Discussion and Conclusion

6.4.1. Summary of findings

The children in this study interacted with the robot in a similar manner to which they did with a human interviewer, although some children appeared to have a clear preference for the robotic interviewer, whilst the other children seemed to prefer the human interviewer. The children with SLCN and autism appeared to speak to KASPAR much more than the human interviewer and this pattern continued with the number of key-words used by the children with SLCN and autism. However, the analysis of the key-points reveals that 4 out of the 5 children revealed less key-points to KASPAR than the human interviewer. One possible reason for this is that the children may have felt less pressure to provide information to the robot than the human interviewer, however this would need further investigation. The results of the engagement coding were consistent with the temporal and transcription data and seem to suggest that the children with SLCN and autism appeared to be more engaged in the interaction with the robot than the human interviewer. Engagement is an important factor in any interaction because the more engaged a child is in an activity, often the better their participation, which in an interview setting could be very beneficial. From a qualitative perspective the children were all very different during their interactions. Some of the children had a very obvious preference for the KASPAR robot, which was reflected in the engagement coding, whilst others responded in a very similar manner in both interviews. It would appear that KASPAR seemed to best suited for the children with communication difficulties and autism.

6.4.2. Relation to literature and advances

Previous work in assistive robotics for children with special needs suggested that some children with special needs such as autism respond well to robots, particularly for therapeutic and educational purposes [3, 9, 71]. The results from this study are consistent with current literature and build on the literature by indicating that some children with special needs respond well to robots in an interview setting which has not previously been tested. Further to this we found that the group of children that using a robot appears to benefit the most is those with communication difficulties. The literature in Section 2.1.5 indicates that children with communication difficulties are often the most difficult children to interview because of their conditions. Considering the results of this study and the literature stating that children with communication difficulties require more assistance in interviews, this avenue of research would likely be worth pursuing in the future as there may be potential for real world impact.

6.4.3. Limitations

This study was a small case study with just five children, had there been more children participating in the study the results may have been clearer. The children in this study were individuals with various different needs and conditions, therefore the results of this study cannot be generalised to a population with particular condition as there is currently insufficient evidence. In addition, although the questions in this study were assessed and approved by the deputy head school teacher, the questions may have benefited from further consideration.

6.4.4. Review of research questions and hypotheses

This study aimed to answer the forth research question: "RQ-4: How do children with special needs respond to a robot compared to a human interviewer?", and provide additional data to answer the first three research questions: "RQ-1: To what degree do children respond to a robot in an interview setting?", "RQ-2: How do children respond to a robot compared to a human interviewer?" and "RQ-3: Will question type and difficulty impact on children's responses to a robot in an interview setting?". The majority of the findings from this study were consistent with the results of the previous two studies, showing that children respond to a robotic interviewer in similar way to which they do a human interviewer, and provide further data to answer the first three research questions. Further to this we found that some of the children in this study showed a clear preference towards the robotic interviewer, which supports the initial hypothesis and provides data to answer the fourth research question. However, the amount of key-points that the children revealed was generally lower, possibly because the children felt less pressured to respond to the robot, which was contrary to the initial hypothesis.

6.4.5. Research questions for next study

The next step in this research was to present the work of the first three studies to experts that work with children and gather feedback on the possibility of a RMI system, to ascertain if such a system would be useful to experts and if so, what features would be required from the system in an effort to answer the fifth and sixth research questions: "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?" and "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". Ascertaining if potential real world users think that RMI systems could be useful for them will assist in determining which application domains should be targeted for this research in the future. Establishing what real world users would require from such a system is also important. Professional interviewers do not generally have expertise in robotics or computer science. For a RMI system to be genuinely useful in the real world, the system would need to be sufficiently flexible and user friendly in order to be used by non-technical operators, therefore gathering the relevant information as to what real world users would require from such a system is essential and is where the efforts of the next study will be focused.

Chapter 7: Gathering expert opinions (Study #4)

7.1. Introduction

The first three studies in this research investigated how robots could potentially be used in an interview setting with both neurotypical children and children with special needs. In these studies the research regarding RMIs focused on researchers directly working with children to test the concept and establish how children respond to a robot in an interview setting. The next logical step is to establish what real world users would require from a RMI system. If robots are to be used to conduct interviews in a real world setting, simply proving that RMIs works in theory with a rigid set of questions and a technically competent operator at the controls is not enough. The system needs to be tested by an expert who has been specifically trained and is experienced in interviewing children to determine if this is a an approach that would really be useful for experts. To address this question we conducted three separate user panels with a variety of potential users groups to gather expert opinions on the first three studies, and what changes would be necessary to make the system suitable for real world users. Feedback on these specific areas allowed us to critique the research conducted in the first three studies and outline a clear set of requirements that real world users would require from a RMI system.

7.1.1. Research questions

The first three studies in this research investigating the potential of RMIs have focused on establishing whether children will respond to robots in an interview setting and if so how well. The results of these studies indicate that both neurotypical children and children with special need will talk to a robot in an interview setting and respond in a similar way to which they talk to a human interviewer. These studies provided us with sufficient data to answer the first four research questions. However, to test if a RMI approach would work in a real world setting, it is important to establish what the experts (specialist working with children) would require from such a system, bringing us to the fifth and sixth research questions: "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?" and "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". To establish the needs of potential users we recruited 20 experts from a variety of backgrounds to participate in three separate user panels. In these panels we gathered the participant's views of the previous studies and ascertained what they would require for the system to be useful to them.

7.1.2. Hypotheses

The previous studies have been very structured in nature to create an accurate and fair comparison between a human and a robotic interviewer. We hypothesised that the main requirements that would be outlined by the participants of the users would be increased flexibility and ease of use. In addition to this we also expected that target group of children that the participants would view this system as most appropriate for would be children with special needs, as this is where the literature in Section 2.1.5 suggests the most assistance is needed.

7.2. Methods

7.2.1. Participants

The three user panels were conducted at separate locations with a total of 20 participants. The panel of specialist police officers police consisted of 11 participants, the intermediaries' panel consisted of 5 participants', and the interdisciplinary panel consisted of 4 participants. Although the police and intermediary panels contained professionals from two specific organisations, there was a mix of participants that performed a variety of different rolls with differing and diverse skill sets. Because of the exploratory nature of this research, the diversity of the panels aided in viewing the potential areas where this technology could possibly assist and also served to create a relatively complete picture of what would be required by the different user groups. The participants of the panels had no prior knowledge of this particular area of research involving robotics. In accordance with the ethics procedure, informed consent was obtained in writing from the participants of the study as outlined in Section 3.1.2.3. The ethics protocol number for this study is 1213/18, the relevant documentation can be seen in Appendix 5.1.

7.2.1.1. User panel #1: Police specialists, 11 participants

The first panel was held with a specialist joint child protection investigation team. The team consists of both police officers and social workers that specialise in working with children. The members of the team continually undergo specialist training that provides them with specific skills to interview vulnerable victims and witnesses. The types of cases that the team would deal with include, internet-based child abuse, complex abuse enquiries within a family environment, sexual abuse of children, and child homicide. In many of these cases it is necessary to interview the children to gather evidence, therefore conducting interviews with children is a routine job for many members of this team. Since the members of this investigation team have to deal with such a wide variety of cases on a daily basis, it puts them in a strong position to provide meaningful feedback about the possibility of a RMI approach in a police interview setting and what would be required from such a system.

7.2.1.2. User panel #2: Intermediaries, 5 participants

The second panel was held with a team of intermediaries that specialise in facilitating communication with children that have various special needs including: learning disabilities, autism, mental health difficulties, ADHD, and physical disabilities. The intermediaries have particularly strong expertise with children who communicate without speech, have complex health care needs, or are affected by other conditions such as autism. The intermediary team includes both adults and young people who have a disability, and advocates who can communicate using sign, symbols and a variety of activity or play based approaches. The expertise of the intermediaries means that they have advocated through many complex situations including: child protection investigations, criminal court proceedings and family court proceedings. The team was founded over fifteen years ago and has consulted with more than 4000 children and young people over that period, and are therefore in a strong position to offer sound advice and feedback.

7.2.1.3. User panel #3: Interdisciplinary user group, 4 participants

The final panel was a multidisciplinary group consisting of, an educational specialist, a magistrate who was previous the head of a service for young people with medical conditions, a paediatric nurse and health visitor, and a specialist in speech and language who has a background in teaching the deaf. The diversity of this group provided a wide variety of perspectives on the potential applications for a RMI system. This panel coupled with the first two panels would serve to give a very clear and relatively complete overall picture as to what potential real word users working with children would require from a RMI system along with the potential advantages and pitfalls associated with such a system.

7.2.2. Procedure

The sessions began by giving a brief background to the research, in particular the previous studies that had taken place. The methodology used for the user panels was a focus group methodology [102]. The background covered during the introduction was an overview of the KASPAR robot, detailing the previous work conducted with children with autism. This was followed by a detailed description of the studies taken place. Details of the studies provided to the participants included: purpose of study, structuring of experiment, topic of interview and findings of the studies. A sample video of the study was shown to the panel immediately after the details of the study were explained to give the participants of the panel a clear example of the RMI system in practice. The videos that were selected were chosen to show the RMIs maximum potential. The selection of videos illustrating work from previous studies allowed us to gather feedback from the panellists and answer the fifth research question: "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?". Immediately after the introduction and background, the participants were asked a number of questions. The principle investigator acted as a low-level moderator when facilitating the discussions to allow the groups to freely explore the domain and options [102]. The questions that the participants were asked are as follows:

UPQ-1) From what you have seen, do you think that this is a tool that could be useful to you? If so…

UPQ-2) Do you think you would use a tool like this? If so…

- How and at what stage?
- With what children?

UPQ-3) Can you think of any specific scenarios where a tool like this would be particularly useful?

UPQ-4) How would you expect to operate this robot?

UPQ-5) What would be the most important features you would expect to see in an interface?

After gathering feedback for these questions the participants were then briefed on the operation of the robot as described in Section 3.2.4. We specifically avoided giving details of how the robot was operated prior to the initial feedback in an effort to avoid constraining the thought process of the participants on how the robot should be operated. After the participants had been fully briefed on how the robot was operated currently, we begun to gather a set of features that the participants would require from a RMI system to answer the sixth research question: "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". The questions that we asked the participants were as follows:

UPQ-6) What features must the interface of the robot have?

UPQ-7) What features would you like to have?

UPQ-8) Could you rank the importance of the features?

These questions concluded the session and provided a clearly defined set of user requirements to adhere to, when developing a new interface that could be used by an expert working with children.

7.2.3. Measures

Whilst the panels were being conducted notes were taken to highlight the key points, in addition to this the sessions were recorded and later transcribed to capture the points and feedback that had not been noted during the sessions[9]. Each panel lasted approximately 90 minutes. The notes and transcriptions of the panels were later analysed and used to generate a list of the key points mentioned in each panel. The lists from each panel were then merged and entered into a table which clearly indicates which key points were stated in each panel. By constructing such a table we could establish which points were universally important and which were more specific to a particular user group, see Table 7.3.3.1.

7.3. Results and discussion

Since there was some repetition in the panels, we first summarise the findings of each panel before going into more depth and consolidating the detailed findings of the specific aspects discussed.

7.3.1. Summary of each user panel

7.3.1.1. User panel #1: Police specialists

The overall summary from this panel was that the participants would only be interested in using a RMI system with children that have special needs, because this is where they believe that most potential for the system lies. The participants stated that as the system currently stands it would not be useable in their field because of the limited and scripted nature of the system. Nevertheless, if the flexibility of the system were to be improved this approach may be useful for some cases involving children with special needs and communication difficulties. The participants did however lend a word of caution about the acceptability of evidence acquired by the robot and stated that this could prove problematic in a court of law, and that this would need to be investigated in more detail before they could rely on such an approach. It is possible that one test case would need to be tried first to establish if this approach could pave the way for a RMI system to be used in future cases, but until this happened it would be an unknown.

7.3.1.2. User panel #2: Intermediaries

The overall verdict from the intermediaries' panel was that the system would only be useful in a small number of cases and could not generally see the system being useful to them in most cases. The participants considered that the system would only be useful for the most disturbed children for who any human interaction is difficult, or for children with conditions such as autism where the child

[9] The first user panel was not recorded due to technical difficulties

finds interaction with people stressful. The participants of the panel stated that in order for this type of system to be a useful tool for them it would need to “up its game” and “do more on its own”. The participants of this panel generally had much higher expectations of technology and would expect the system to be autonomous, quoting systems such as Apple’s Siri. This expectation may be caused by media hype and a misinterpretation of current voice analysis systems. Because a system will listen, process a command and respond, this does not mean that it has a true developed cognitive ability to understand, therefore such systems would not be suitable for such applications. The participants stated that for the system to be useful in some form, the lag time would need to be reduced and the flexibility increased, similar to the police panel participants.

7.3.1.3. User panel #3: Interdisciplinary user group

The participants of the interdisciplinary panel shared the views of the previous two panels in the belief that the flexibility of the system would need to be improved in order to make the system useful in a real world setting. The participants generally had a very positive view of this approach and could see a number of settings where this type of system could be used, including the health system for child protection and a medical context for asking children where it hurts, as sometimes it is difficult to get a child to respond. There were also suggestions of the system appealing to children’s mental health workers as they may find the system useful for counselling. One of the key areas that this panel highlighted in particular was the importance of usability, reliability and the need for training. The participants explained that guidelines would need to be developed to establish which children this approach would be best suited to (i.e. age range), and that a training program would need to be developed, showing users how to fully utilise the system. Although the participants considered that an effort should be made to establish which children this system would be best suited to, they did state that in some cases it would be difficult to tell how a child will respond until you have actually tried it. The most important aspect of the system from the participant’s perspectives was the reliability of the system, stating that the system should be robust and it must not break down, as the professional using the system needs to have confidence in the system. In addition to this the participants emphasised the importance of having a system that is easy to operate, stating that if it’s difficult to use people won’t want to use it. Overall the response from this panel was very positive and the participants could see the potential in a wide range of application areas. Because this panel consisted of individuals from a variety of different professional backgrounds, this may have increased their ability to visualise this system in a wider range of application areas.

7.3.2. Detailed findings

The findings from the panels are split into 5 sections each of which addresses an aspect relating either to the three studies conducted or suggestions for the future direction of the work, see Table 7.3.2.1.

Table 7.3.2.1: Each section answers one or more of the User Panel Questions (UPQ)

Section	UPQ
Feedback on previous work	1
Would experts use this system and where do the main benefits lie?	1, 2, 3
Potential complications and advantages of Robot-Mediated Interviews	1, 3
Aesthetics of the robot	4
Interface and operation of the robot	4, 5, 6, 7, 8

7.3.2.1. Feedback on previous work

When requesting feedback on the previous work, we found that the participants from the first panel of police officers and social service staff were not surprised that the children spoke to the robot in a similar manner to which they did the human interviewer. The participants explained that children will often talk to puppets in a similar way to which they talk to a person, even if it is obvious that the puppet is being controlled by someone else. To support this comment additional research has shown that puppets have been used to help alleviate stress in child undergoing medical procedures [103, 104]. The participants of this panel also noted that the children in the previous interviews had no stresses associated to the questions that they were being asked, stating that the results may have been different if the questions were of a sensitive or stressful nature. This is a fact that we were aware of, however this type of study would need to be conducted by a professional interviewer that has the appropriate experience and expertise in this field, for ethical reasons we would not be able to conduct this type of study ourselves.

Another aspect associated to the previous work that was discussed, related to visible levels of stress from the children being interviewed. The participants of the police panel explained that in court cases, some level of stress or emotion is expected to be seen from the child in order for the case to be taken seriously by the prosecution. If a child appears to be too comfortable and not showing any signs of emotion or stress, this can actually harm the prosecution, because people often have preconceived ideas about how a child should be reacting when being questioned about sensitive issues. This is a factor that we had not previously considered, however this is only likely to become an issue much further in the future, if evidence that has been collected by a robotic interviewer is used in court proceedings and is therefore outside of the scope of this thesis. The participants of the panel went on to say that a disclosure from a child to a robot that could not be used as evidence in court would still be better than no disclosure at all, because if the child is in any immediate danger measures could be put in place to protect that child.

In all three of the panels there were questions about measuring the consistency of information provided by the children in the studies. One participant from the second panel of intermediaries commented that the children in the videos appeared to be more worried about getting the answers right with the adult interviewer than the robotic interviewer. We explained that the consistency of the information that the children provided in both interviews was measured and that no considerable differences were found in the information that the children provided.

During the second panel with the intermediaries, we were asked if the eye gaze of the interviewer was controlled or analysed, to which we informed them that these aspects had not been measured. The participants explained that adult eye gaze is a control mechanism and that their staff are trained how to use their eye gaze when working with children. The intermediaries explained that using eye gaze appropriately when working with children can make the children much more comfortable. The intermediaries stated that this may account for why the children looked at the robot more than the human interviewer. In addition to this the intermediaries directed us towards literature "don't look now I'm trying to thinking" [105] that suggests children find it easier to think if adults are not looking at them.

The participants of both the second and third panels commented on some of the children appearing to be more comfortable with the robot than the human interviewer. A participant of the third panel noted that there was a "massive difference" in the behaviour of the girl in the third study between the robot and the human, stating that the girl seemed much more relaxed and interested with the robot, and that "the difference is tangible". A participant of the second panel stated that the girl in the video appeared to be more open with the robot than the human interviewer, and that her head is tilted down more whilst talking to the human interviewer. Participants from both panels suggested that a qualitative analysis approach would be better suited for this research, and that perhaps a structured observation would be the best means of capturing the information, as much of the research in this thesis has been based on quantitative results.

An aspect of the previous studies commented on by the second and third panels were the suitability of the interview arrangement. Both panels of participants believed that the interview arrangement would have been better if the child and interviewer were not directly facing each other, stating that it may be better to sit slightly at an angle so eye contact is not forced. Several examples of this were given in both of the sessions. The paediatric nurse participating in the third panel stated that when working with older children it is sometimes better to sit near the child and give them gentle encouragement to talk rather than directly opposite the child, because sitting face to face can be seen as very confrontational and possible a bit scary for the child. She also noted that in a doctor's surgery the patient is sitting side to side with the doctor rather than face to face, and patient comfort is possibly a reason for this. Another example given by the paediatric nurse was from personal experience of working with children who missed education because of mental health problems. She stated that sometimes the best conversations with the children that she was working with were in the car where there is no eye contact, and that this lack of eye contact may have been a contributing factor. The participants from the intermediary panel stated that they would never attempt to talk to a child with special needs face on, and that they would usually sit side on. The participants of the second panel believed that it would be better to have a comparison of the robot and a skilled professional conducting the interviews with children particularly in the third study and to have a side by side arrangement rather than a face on arrangement with the child.

The intermediaries explained that the criminal justice system tries to establish if the child is a competent witness, however the intermediaries argue that it is impossible to answer that question unless the competence of the adult conducting the interview is known. The intermediaries explained that the competence does not exist in the child, it exists as an exchange. For example, if in interview with a 2 year old is conducted well, the evidence that can be obtained will be very good. In contrast if an interview with a 10 year old is conducted poorly, the evidence obtained will be very poor. The intermediaries stated that a RMI approach may be useful if the child's competence is dependent upon an adults' certain kind of behaviour, i.e. not appearing to show any emotions or visually reacting to what is being said.

7.3.2.2. Would experts use this system and where do the main benefits lie?

The participants of all three panels considered that the system would not be useful as it currently stands because it is too restrictive. Each panel clearly stated that the system would need to be far more flexible and have the ability to respond to any question because the nature of the interviews with children is often unpredictable. The participants stated that if the flexibility of the system were to be improved, using the system could be considered. The participants of the police panel in particular considered that the system was unlikely to be used with typically developing children because the techniques used with typically developing children are well established, therefore efforts should be focused on working with children that have special needs such as autism, as these are the children that professionals would be interested in using the system with. Further to this, the general consensus of opinion shared by all three of the panels was that the majority of the potential benefits of a RMI approach would lie with children that have special needs. There are very well documented approaches and successful methods for interviewing typically developing children [7, 18, 106, 107]. However, interviewing children with special needs is much more difficult and as a result prosecution and conviction rates for cases involving a witness with special needs is much lower [11, 12], despite children with special needs being up to 4 times more likely to be a victim of abuse [5]. These facts coupled with the research suggesting that children with special needs respond well to robots, presents a strong case for focusing future investigations in this area on children with special needs.

The participants of the intermediaries' panel had a more refined view of the system and believed that the system would have a relatively small niche being suitable for children with autism or the most disturbed children for whom any human interaction is difficult. For example, KASPAR may be particularly useful when interviewing children that have been sexually abused because it is not the same as talking to a person, and some of the most traumatised children may reveal more to a robot than a human. The intermediaries' generally had a much higher expectation of technology and stated that an autonomous system would be much more useful to them, quoting systems such as Siri on the Apple iPhone. Currently such systems would not be suitable for conducting interviews of a sensitive nature with children. Interviewing children about sensitive or emotionally provocative events is a demanding task that requires specialist knowledge and skills, along with a thorough understanding of the relevant background to a case. Computers currently lack the cognitive abilities to perform such tasks and are therefore unsuitable for this task in an autonomous role.

The healthcare expert in the interdisciplinary panel stated that the system may be useful in a children's hospice to talk to the children about their illness. She explained that the system would have an advantage over a human because it would not appear to get upset when talking about something distressing. In this context the robot would be used over a sustained period to develop a relationship with the child, and further studies would need to be conducted to establish how effective such a robot would be able to fulfil such a role, but it may provide the child with a useful outlet to talk about their illnesses. An example that the participant quoted was of a child that was terminally ill and knew this but did not want to talk about her illness because of the emotional upset it cause her parents. There were very few forums where the girl would be able to talk about her illness and a robot such as KASPAR could act as a useful intermediary for this type of situation because the robot would not get upset. This potential application was something that had not been considered previously, but upon receiving this suggestion it would seem logical that KASPAR may have some potential in this type of role.

7.3.2.3. Potential complications and advantages of Robot-Mediated Interviews

As with any system or approach there are potential advantages and pitfalls associated to using them. Aside from the potential advantages a RMI approach may possess which have previously been stated in this thesis, there are some specific advantages and pitfalls to using a robot in an interview and accompanying reasons why. One of the most obvious advantages that all three of the panels stated was the robots ability to be calm, predictable and simple in its responses. KASPAR won't have preconceptions and could listen with an open mind, not emitting social queues or atmosphere. The participants of the intermediaries' panel in particular noted that the robots lack of complex body language and ability to "not mind" may be an advantage, as that is something that interviewers find very hard. KASPAR could be good at "not getting upset", and not emit the small signs of anxiety or distress which some children detect. The participants stated that children often notice if someone is really shocked or not and they will often detect if someone is faking or trying to suppress their feelings. All three of the panels considered that the robots lack of body language and emotion could be both an advantage and a disadvantage, depending on the child because each child is different.

Another potential advantage mentioned by the interdisciplinary panel was that the mere fact KASPAR is not human, the children may not expect the robot to have expectations or answers. The intermediaries stated that, children sometimes think that the person asking them questions already have the answer in their head and that there is a right and wrong answer. Building on this, the participants of the police panel raised a point about the risk of the children slipping into a fantasy type scenario when using interview aids. The participants explained that one of the difficulties sometimes experienced whilst using props or aids, is that it can encourage the children to treat the interview like a playtime fantasy type scenario. In a fantasy world it would be acceptable for the children to fabricate and fantasise which would obviously present a problem when a child is giving evidence. KASPAR could potentially negate this problem because it has a human like appearance, however this would need further consideration and investigation to ensure that children do not see talking to KASPAR as a fantasy situation. Talking to KASPAR would need to be just as serious and as real as talking to a person. The point raised by the police panel would also indicate that KASPAR may potentially be a more appropriate robot than a robot such as a NAO. Because KASPAR has a human type appearance the children may be less likely to slip into a fantasy type scenario, however this would require further investigation. The results of quantitative and qualitative analysis of the three separate studies with fifty children, indicate that the children interacted with the robot in a similar manner to which they did a human interviewer in terms of the information they provided.

The participants of the police panel suggested a robotic interviewer may have a potential unforeseen advantage with regards to performing interviews in a more fluid uninterrupted form. According the participants of the police panel, currently as the UK rules stand, if a member of the investigation team thinks of a useful question to ask the child while an interview is taking place (even the technical camera person), the child must be asked that question to give them the opportunity to disclose. The participants said that this sometimes results in pieces of paper being slid under the door of the interview room by other officers that are monitoring the interview, which can be very distracting to the child and can sometimes interrupt the interview. KASPAR could negate this problem as a team of officers could control the robot remotely and would be able to offer the child a seamless interview with no such interruptions.

7.3.2.4. Aesthetics of the robot

Aspects of the feedback from the panels related to the visual and audible aesthetics of the system. It was noted by the police panel that the robot looks disproportionate and strangely clothed. The participants of the police panel stated that KASPAR appears to have a full sized male head on a child's body and this could be interpreted as an adult trying to act like a younger person. The participants explained that this could present a problem because individuals that abuse children often try to present themselves in a childlike manner. The participants of the panel felt that the neither adult nor child look of the robot could be a confusing to a child, or actively problematic if the child had been abused by an adult trying to portray the image of a child. This is an important aspect that will need further consideration when building and dressing future implementations of the robot. However, it is also important to take into account children's view of the robot. The KASPAR robot has been used with over 400 children since 2006 and this has not presented a problem in the past. It is possible that the participants of this panel had some preconceptions about how children would respond to a robot with such an appearance. More recent versions of KASPAR are more proportionate, partially addressing this concern. The concerns of an adult trying to impersonate a child were also extended to include the aesthetics of KASPAR's voice. The participants commented that KASPAR's voice was not very childlike and it would be better to have a more childlike voice that was slightly robotic to help maintain the impression of the robot. Despite the points raised by the participants that need attention, the general consensus of opinion about having a robot with a simple neutral face that could be controlled not to show emotion was positive. All of the feedback gathered relating to areas for improvement are addressable and will be considered in future implementations of the robot. The key principal of a humanoid robot with a human like face interviewing children seemed to be embraced by the majority of the participants.

7.3.2.5. Interface and operation of the robot

The overall consensus of opinion from all three panels of what would be required for a RMI system to be useful in a real world setting was very similar. Currently as the interface stands it was universally agreed that the system would not be useful to experts working with children because of the lack of flexibility. All participants maintained that the system would need to be more flexible, with freedom to respond to and ask questions spontaneously rather than in a scripted manner similar to the first three studies. Interviews with children, particularly police interviews, are often very spontaneous in nature and the direction of the interviews is often unknown and unpredictable. A participant from the police panel stated: "no two interviews are the same, you try what you think will work and if it doesn't you will use your experience and try another approach". All three panels came to the same conclusion that the most effective means of increasing the flexibility would be to have a direct link to the robot where the interviewer's voice would be converted and spoken by the robot in the remote location. The participants of the intermediaries' panel were more interested in having an automated system. However, when the limitations of current technology and the system were explained to them, they felt that a direct voice conversion option would be the next best option. With regards to automation, the participants of the three panels believed that some automation would be useful for behaviours such as blinking, so that the interviewer could focus on interviewing the child rather than controlling the robot. The participants of the panels indicated that body language is less important and should be kept to a minimum. In all of the panels it was commented that too much body language could be distracting to the child. In addition to this, research suggests that inappropriate body language can actively mislead witnesses in interviews, see literature in Section 2.1.4. Having a robot whose gestures can explicitly be controlled could remove this problem all together. The three panels also stated that the system must be reliable, because an unreliable system would cause the interviewer to lose confidence and adversely affect any interview. The list of features/specifications that the panels felt were most important can be seen in Table 7.3.2.5.1.

Table 7.3.2.5.1: The user panels ranked the importance of the features they would require from a Robot-Mediated Interviewing system and stated the reason why

Priority	Feature	Reasoning
1	Reliability	The system must be reliable and work every time. If it is not reliable, everything else is insignificant.
2	Flexibility	The system needs to be flexible to follow any avenue of questioning as the path of questioning is often unknown.
3	Usability	The system must be easy to use and allow the interviewer to focus on interviewing the child rather than operating the system. Some level of automation may need to be employed to achieve this.
4	Minimal body language	The system should employ a minimalistic use of body language as this may distract the children.

7.3.3. Summary of user requirements

The comments and requirements of the three panels are listed in Table 7.3.3.1.

Table 7.3.3.1: Each user panel generated opinions and suggestions which are summarised in this table			
Comments	**Panel comment matrix**		
Feedback on previous work	**Panel #1**	**Panel #2**	**Panel #3**
Not surprised that children responded to robot	X	X	
Some emotional distress expected from child in order for the case to be taken seriously	X		
Evidence that cannot be used in court is better than no evidence at all	X		
Was the information that the children provided consistent with the robot and the human?	X	X	X
Children in videos appear to be more concerned about getting the answers right with the adult than the robot		X	
Because the eye gaze of the human interviewer was not controlled for, this may have impacted on the results of the interviews in the three studies		X	
Children appear to be more comfortable with the robot than the human interviewer, particularly one case in the third study		X	X
Qualitative analysis may be better suited for this research		X	X
Interviews may have been better if the child and the interviewer were not directly facing each other, so eye contact is not forced		X	X
The criminal justice system try's to establish if the child is a competent witness, however the competence does not exist in the child, it exists as an exchange. A RMI approach could be useful if the child's competence is dependent upon an adults behaviour		X	

Would experts use this system and where do the main benefits lie?	**Panel #1**	**Panel #2**	**Panel #3**
System not useful as it currently stands because it is too restrictive	X	X	X
The system would need to be far more flexible and have the ability to respond to any question because the nature of the interviews with children is often unpredictable	X	X	X
It is unlikely that such a system would be used with typically developing children	X	X	
The majority of the potential benefits may lie with children that have special needs		X	X
The system would have a relatively small niche, useful for children with autism or the most disturbed children for whom any human interaction is difficult		X	
The system may be particularly useful when interviewing children that have been sexually abused because it is not the same as talking to a person		X	
Much higher expectation of technology, stating that an autonomous system would be much more useful, quoting systems such as Siri on the Apple iPhone		X	
The system may be useful in a children's hospice to talk to the children about their illness			X

Potential complications and advantages of Robot-Mediated Interviews	**Panel #1**	**Panel #2**	**Panel #3**
The majority of the potential benefits may lie with children that have special needs	X		
The robots lack of body language and emotion could be both an advantage and a disadvantage, depending on the child	X	X	X
Potential risk of child slipping into a fantasy type scenario with a robot	X		
Potential advantage of performing interviews in a more fluid uninterrupted form because multiple members of the investigation team can work behind the robot and ask different questions from one interface	X		
The robots ability to be calm, predictable and simple in its responses could be advantageous		X	X
KASPAR won't have preconceptions and could listen with an open mind, not emitting social queues or atmosphere which could be useful		X	
Children sometimes think that the person asking them questions already have the answer in their head and that there is a right and wrong answer. The mere fact KASPAR is not human, the children may not expect it to have expectations or answers			X

Aesthetics of the robot	Panel #1	Panel #2	Panel #3
KASPAR appears to have a full sized male head on a child's body and this could be interpreted as an adult trying to act like a younger person. This could present a problem because individuals that abuse children often try to present themselves in a childlike manner. The neither adult nor child look of the robot could be confusing to a child, or actively problematic if the child had been abused by an adult trying to portray the image of a child.	X		
KASPAR's voice is not very childlike and it would be better to have a more childlike voice that is slightly robotic to help maintain the impression of the robot	X	X	X
Interface and operation of the robot	**Panel #1**	**Panel #2**	**Panel #3**
Currently the system would not be useful to expert interviewers due to the limitations and lack of flexibility	X	X	X
The system needs to be more flexible with freedom to respond to and ask questions spontaneously	X	X	X
The most effective way of increasing flexibility would be to have a direct link to the robot where the interviewer's voice would be converted and spoken by the robot in the remote location	X	X	X
Body language is less important and should be kept to a minimum because too much body language may be distracting to the child	X	X	X
Some degree of autonomy may be useful to allow the interviewer to focus on interviewing the child	X	X	X
The most important feature of the system must be reliability	X	X	X
Appropriate guidelines and training would need to be established in order to ensure the professional is getting the maximum benefit from the system			X

7.4. Conclusions and direction for future interface development

7.4.1. Summary of findings

The panels that we conducted had many different tangents of ideas, but there was an overall general consensus of opinion from all three panels about what features the system must possess and what children the system would most likely be useful for. All three panels shared the view that the flexibility of the system would need to be improved in order to make this system useful in a real world setting. The participants of the first and last panel generally had a very positive view of this approach and could see a number of settings where this system could be used, including the health system for child protection and a medical context for asking children where it hurts, as it is sometimes difficult to get children to respond. The common view shared by all of the panels was that the system would be best suited to working with children that have special needs. The participants in the police panel in particular stated that they would only be interested in using such a system with children that have special needs because they believed that this is where the greatest potential for the system lies and this is often where addition measures need to be taken to help the child. This is consistent with the guidance of the ABE which states “It is important to find out what impact the child’s condition is likely to have on the interview or on the communication process, and to adopt a positive approach that focuses on the child’s abilities when trying to find out how they can be helped to communicate.” [7],p.172. In addition to this there were also suggestions by the participants of the third panel that the system may appeal to children’s mental health workers or clinical psychologists, as they may find the system useful for counselling.

The panels all stated that as the system currently stands, it would not be very useful to them in a real world setting due to the limited scripted nature of the system. However, the panels stated that if the flexibility of the system were to be improved, this approach may be particularly useful in cases involving children with special needs and communication difficulties. The participants of the police panel did however express doubts about the legal acceptability of evidence acquired by the robot and stated that this could prove problematic in a court of law, and that this would need to be investigated in more detail before they could rely on such an approach. It is possible that one test case would need to be tried first to establish whether this approach could be used to pave the way for future legal cases.

One of the key areas that the participants of the third panel emphasised in particular was the importance of usability, reliability and the need for training. The participants explained that guidelines would need to be developed to establish which children this approach would be best suited to (i.e. age range). It was also stated that a training program would need to be developed to show users how to fully utilise the system. The participants stated that in some cases it would be difficult to tell how a child will respond until you have tried it. The most important aspect of the system from the participant's perspectives was the reliability of the system, stating that the system should be robust, as the professionals using the system would need to have confidence in the system and that it should be easy to operate.

7.4.2. Relation to literature and advances

Some research on child-robot interactions in sensitive contexts has investigated the needs of the users from a variety of perspectives [54]. However, to date all studies focusing on the potential application area of RMIs have focused on the interactions between the interviewer and interviewee without an investigation of what requirements real world users would need from such a system. This study has approached the application area from a different angle and has established what professionals using such a system would require in order for the system to be useful for them. Asserting what users would require from a system is an important step in creating a system that could be used in the real world rather than just in a research context.

7.4.3. Limitations

The feedback gathered in the panels was useful for outlining a set of requirements for a RMI system, however it would have been useful to complete several cycles of feedback, implementation and testing to compile a thorough set of requirements. Physical testing of a system always highlights issues that sometimes go unnoticed in a theoretical walk through of a system. In addition to this it would have been desirable to perform a thematic analysis [108] of the data collected from the user panels.

7.4.4. Review of research questions and hypotheses

The feedback gathered from all three panels indicated that a RMI approach would be considered by various professionals provided certain requirements were met which answers the fifth and sixth research question: "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?" and "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". In terms of the hypotheses, as expected two of most important features to the potential users of the system were flexibility and ease of use, however additional criteria were also outlined. The requirements outlined are as follows:

- The system must be robust and reliable, so that the professional has confidence in the system
- A direct speech interface that converts the voice of the interviewer allowing them sufficient flexibility to respond to any unexpected and spontaneous path that an interview may take.
- A small selection of gestures that can be activated by a visual Graphical User Interface, whether this is in the form of a touchscreen or mouse based device.
- Autonomous behaviours such as blinking to enable the interviewer to focus on the task of interviewing the child rather than operating the robot.
- A sufficient training program would need to be implemented in order to ensure that the professionals using the system could maximize its potential.

With regards to which children, the participants felt the system would be best suited to children with special needs as hypothesised. In addition to this there were a number of other suggestions that had not been considered, in particular for children with serious illnesses who find it difficult to talk to people about their illness because of the upset that it causes people. This is an application area that had not previously been considered but would possibly be worth investigation in the future.

7.4.5. Research questions for next study

The findings from this study have generated a relatively clear set of requirements. The next phase of this research focuses on implementing these requirements, then testing the system with a potential real world user to gather their views on such an approach. Assessing the system with a potential real world user aims to answer the final research question: "RQ-7: Could robots such as KASPAR be used in an interview setting effectively by an expert working with children?". Testing a RMI approach with a potential user is one of the most important steps in establishing if such a system would work in a real world situation and genuinely benefit experts working with children.

Chapter 8: Field trial with a potential real world user (Study #5)

8.1. Introduction

The first three studies conducted demonstrated that children respond to a robot in a similar manner to which they do a human interviewer regardless of question difficulty. The fourth study investigated how an expert working with children perceived a RMI system and whether this approach may be useful in their field. The feedback gathered from the 20 participants in three panels indicated that a RMI approach might be useful in a number of areas provided that the system was sufficiently flexible. The next step in this research was to trial the system with a potential real world user. Prior to conducting the field trial, modifications were made to the KASPAR program in accordance with the feedback obtained from the user panels to ensure that the system would be sufficiently functional for a real world user. The study involved a pre-trial introduction and interview with the expert trialling the system to acquire a detailed description of their professional background and experience, along with their expectations of the system being trialled and give an overview of the arrangements for the study. The pre-trial was followed by the trial in a school where the expert interviewed 10 children in a mainstream school individually regarding a 3 minute video they had watched prior to the interview. The study then concluded with a post-trial interview with the expert to gather feedback on the system. Based on results reported in chapter 6, ideally we would have recruited children with special needs for this study. However, after the system modifications had been made, based on feedback from experts presented in chapter 7, the final study focussed on the usability of the system by an expert user, rather than on the children's interview responses. We therefore recruited an opportunity sample from a local school.

8.1.1. Research questions

The study described in this chapter aims to answer the final research question, RQ-7: Could robots such as KASPAR be used in an interview setting effectively by an expert working with children? and provide additional material to answer the fifth and sixth research questions, "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?" and "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". Trialling the system with a potential real world user is a critical step in establishing if this approach could be feasible outside of a research context and ultimately have real world implications.

8.1.2. Hypotheses

Three consecutive studies have shown that both neurotypical children and children with special needs respond to KASPAR in an interview setting. Coupled with the modifications implemented to make the system more flexible as outlined by the user panels, we hypothesise that system will perform successfully. However, we expect to gather additional requirements because actual use of the system is likely to highlight unforeseen issues.

8.2. Methodology

8.2.1. Technical modifications

Prior to commencing the study, the feedback obtained from the user panels was considered and the most important elements were implemented, these were:

- A Direct microphone link to the robot with voice conversion to increase the flexibility of the system.
- Automated behaviours to allow the interviewer to focus on the child rather than controlling the robot.
- A selection of gestures that could be activated by the onscreen GUI.
- Based on the previous studies the system was already relatively robust and therefore no additional modifications were required for this aspect of the system.

8.2.1.1. Voice conversation system

The participants of the user panels stated that the system would need to be more flexible and easy to use. Therefore a direct microphone link with some voice conversion was implemented. The software used to implement the voice conversion was Screaming Bee MorphVOX Pro 4.4.8 [109]. The software allowed for detailed customisation of the voice projected, which suited the needs of this application. For the purpose of this study we required the voice of the interviewer to sound childlike, robotic, but most importantly, understandable. In previous studies some of the children had difficulty understanding the robots voice, which had been generated from text to speech software. Taking this into consideration, the customisation of the voice output was very important and needed to be constructed carefully. To strike the correct balance many different settings were tried and tested in the software. When testing the different settings, a non-native English speaker was present to give his feedback on the clarity of the voice. This was important, because understanding a voice with an accent is often more difficult for non-native speakers, therefore making him an ideal candidate to test the understandability of the voice. The settings that were finally chosen increased the pitch of the voice slightly to make it sound more childlike, and a slight de-tune effect was implemented to create a more robotic tone and reinforce the impression of the robot. To thoroughly test the clarity of the voice, the principle investigator spoke to the tester in a remote location with the voice conversion activated and asked him to repeat what was being said. By testing the voice in this manner it was possible to provide an accurate indication of how easy the voice was to understand.

8.2.1.2. Automated behaviours

One of the requirements outlined by the user panels referred to ease of use of the system. One participant in particular stated that the system should allow the interviewer to focus on interviewing the child, rather than operating the system. In order to do this the blinking and mouth movements of the robot were automated. The blinking of the robot was generated by randomly within a set threshold by a timer. In addition to this, automated mouth movements for the robot were used to maintain the impression that the robot was talking whilst the interviewer was asking questions via the robot. The code used to achieve the mouth movement implemented a simple threshold mechanism. The openness of the robots mouth would increase according to how loud the interviewer spoke, if the interviewer was not speaking the mouth would be closed. It is important to note that we were not looking to create lip-sync. The KASPAR robot would not be capable of producing perfect lip-sync due to the limitations of the hardware, in addition to this, creating ultra-realistic lip-sync was not necessary to test the concept, which was the purpose of this study. The code for these automated behaviours can be found in Appendix 6.2 and 6.3. Whist creating the two primary automated behaviours, two additional behaviours were also developed, eye noise and body noise. Because people make slight involuntary movements, code was developed and tested to make the robot fidget very slightly and to create very slight eye movements. These pieces of code were not activated in the final system that was trialled with the expert, because it was felt that these small movements may be distracting to the children, however, the code for these behaviours can be found in Appendix 6.4 and 6.5.

8.2.1.3. Gestures implemented

A selection of gestures were implemented in the software, in accordance with the feedback gathered from the user panels. The gestures were activated via a Graphical User Interface, which can be seen in Figure 8.2.1.4.1. A selection of the robots gestures can be seen in Figure 8.2.1.4.2.

8.2.1.4. Overall installation

The overall study setup was similar to the first three studies conducted in terms of having a remote connection to the robot and a camera to monitor the situation remotely. However, in this study the modifications to the system were implemented as described in the sections above and the situation was monitored by a webcam which was connected to the PC that controlled the robot. The room layout can be seen in Figure 8.2.1.4.3.

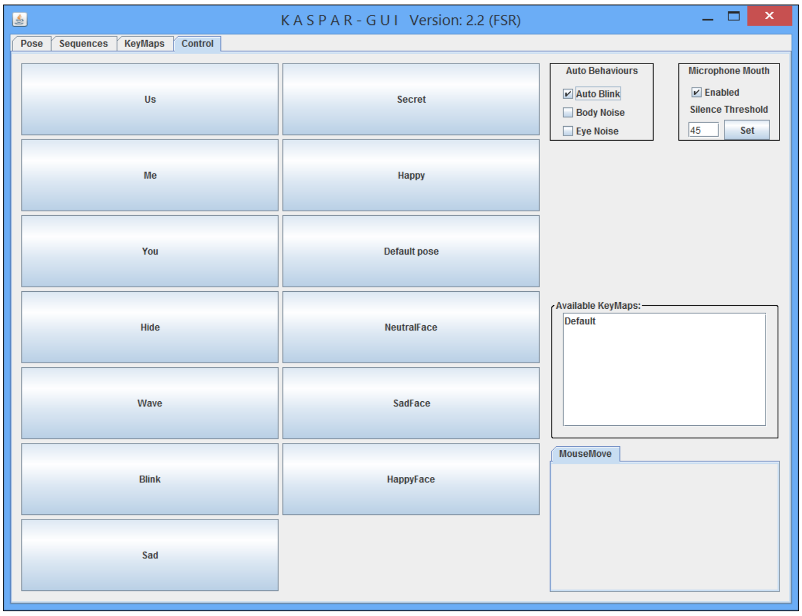


Figure 8.2.1.4.1: Screenshot of modified KASPAR Graphical User Interface (GUI). Note: additional features in top right hand corner.

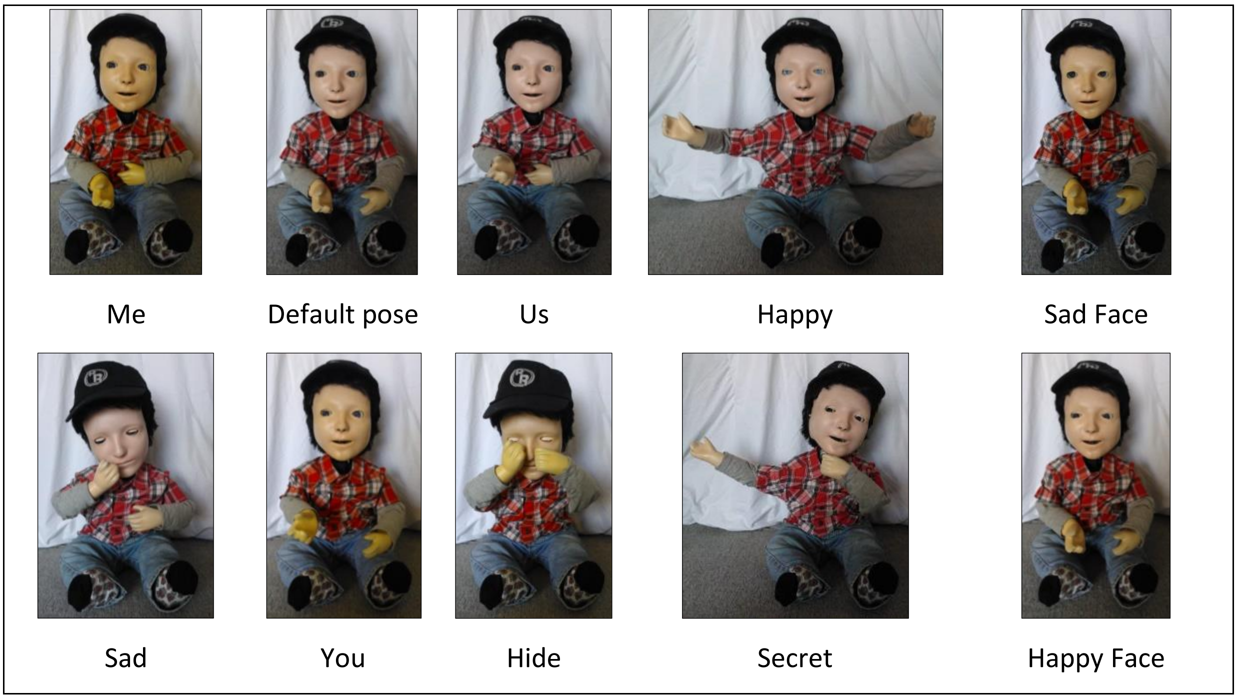


Figure 8.2.1.4.2: A selection of the robots gestures

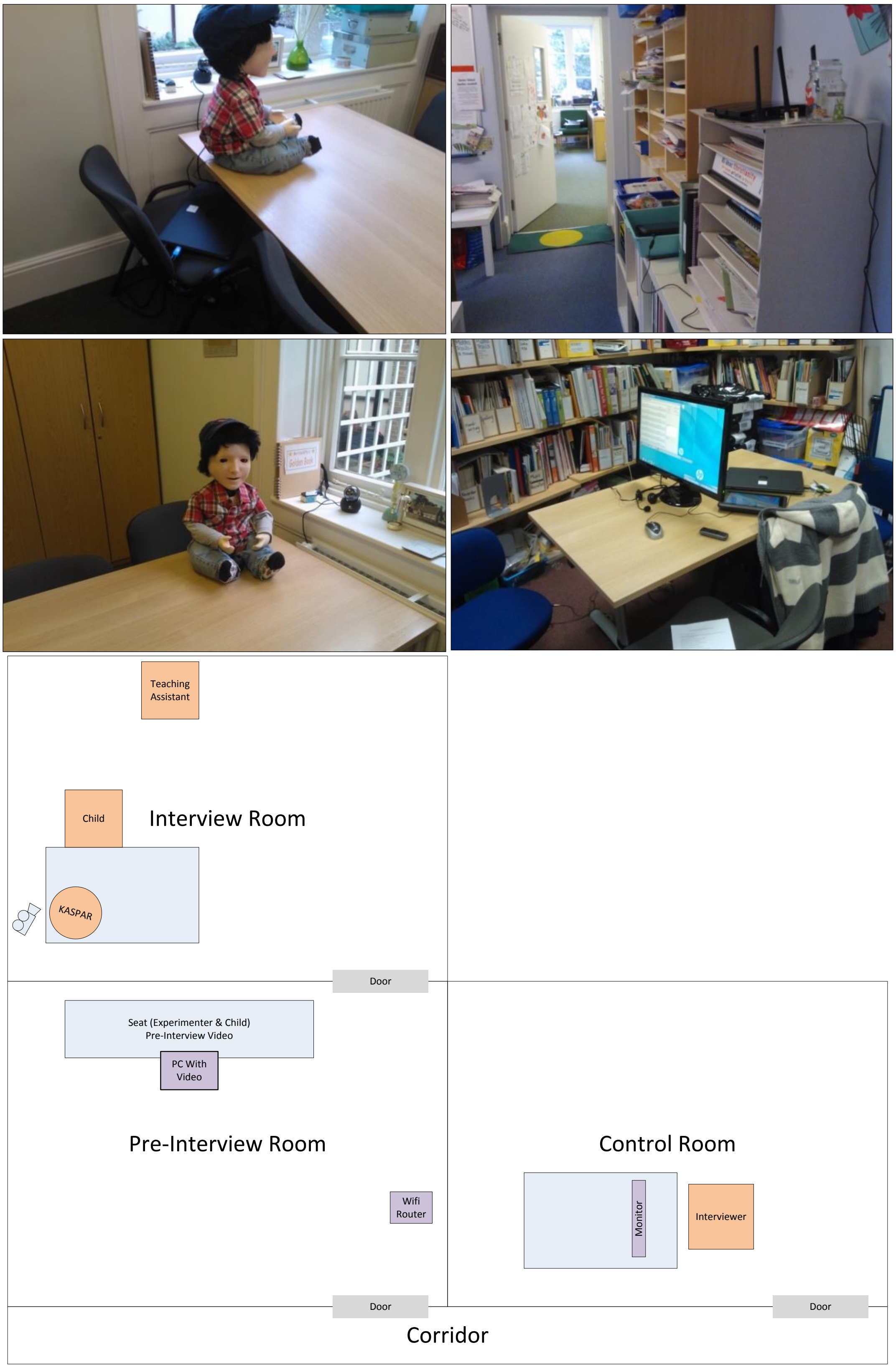


Figure 8.2.1.4.3: An overview of the room layout and interview setup

8.2.2. Participants

This study included one adult and ten child participants. The adult participant was a senior Educational Psychologist with over 20 years of experience. The child participants were ten neurotypical children between the ages of 7 and 8 (6 male and 4 female). The study was conducted at a local mainstream primary school in Hertfordshire in the UK. This age group of children was chosen on the basis of the previous research we conducted. In order to provide the expert with a wide spectrum of interviews the class teacher was asked to select a variety of children with different personalities, according to her judgment. In accordance with the procedure outlined in Section 3.1.2.3, informed consent was obtained in writing from all parents of the children participating in the study and the adult interviewer, agreeing to take part in the study. In addition to this, written confirmation was also obtained from the Head of the school to confirm that she was happy for the study to take place in the school and acknowledged that she had seen and was satisfied with the 3 minute video clip that the children were to be shown and questioned about. The ethics protocol number for this study is COM/PG/UH/00013, the relevant documentation can be seen in Appendix 6.1.

8.2.3. Procedure

The study was conducted in 3 phases for both the children and the adult interviewer, utilising a pre-trial introduction, trial run, and post-trial debrief.

8.2.3.1. Pre-trial introduction (the children):

The children were briefly introduced to KASPAR approximately 3 weeks before the trial commenced. This introduction served to familiarise the children with the robot and was also used to explain that KASPAR would be coming in to talk to some of them. In addition to the KASPAR robot, the children were also given the opportunity to play with some other robotic toys because the children's school term was going to be themed on robots, therefore the class teacher considered that this would be a positive start to the term for the children.

8.2.3.2. Pre-trial introduction (the interviewer):

The professional interviewer was given a pre-trial introduction to explain the procedure for the trial. It was explained that all of the children had been briefly introduced to KASPAR as a group to familiarise them with the robot, and that the children were aware that KASPAR would be coming back to talk to them. The interviewer was advised that the children would be shown a 3 minute video clip that KASPAR would then ask them questions about. The interviewer was given a minimal description of the video to help with potential lines of questioning, which was as follows:

"The cartoon depicts two animal characters sitting down in a house with a radio playing in the background. All is quiet and normal before some creatures come into the house and cause trouble by eating various items. The two characters eventually combine their efforts to get rid of the creatures that are causing the trouble."

The interviewer was specifically not shown the video prior to the interviews because we wanted to create a situation similar to a real-life interview concerning an event that the child had witnessed, but the interviewer did not. It was explained that once all of the interviews had taken place, the interviewer would get to see the video and would then be able to establish how much information he had managed to elicit from the children. In this pre-trial introduction it was also stated that questioning must only focus on the contents of the video that the children had seen as this was in accordance with the ethics clearance. Following the procedural explanation and demonstration of the system, an interview took place with the psychologist to get a detailed description of his professional background and experience. The pre-trial interview questions were designed to establish: what the role involved, how much experience he had in the role, what the most important aspects for the role were, how technology currently plays a part in that role and what expectations he would have of a RMI system. The questions asked were as follows:

- Job title:
- Number of years in role:
- Can you describe the duties of your position/role?
- What are the most difficult aspects of your profession?
- Describe two or three major trends in your profession today.
- What are the characteristics that are most important for your role?
- Describe your working environment and the role that technology plays in your day-to-day job.
- How familiar are you with electronic consumer devices? What devices do you use and how often are you using them?
- How often do you carry out interviews questioning with children?
- When communicating with children what causes the most difficulties?
- How do you overcome these difficulties?
- What are your expectations of a robot that you can use as an interface for interviewing children?
- What preconceptions do you have about the advantages and limitations of a Robot-Mediated approach to interviewing children?

8.2.3.3. The trial (the interviewer and children):

The trial took place 3 days after the pre-trial interviewer introduction. The interviewer was given a brief reminder of the procedure and a short test of the setup prior to the interviews with the children commencing. The school had received 16 parental consent forms, and of the 16 forms, 10 were chosen by the teacher to take part. The teacher was asked to select as wide range of different personality types as possible, as this would give a good overall representation to the interviewer as to how different children would respond to the robot. The children were taken to and from their class by a teaching assistant. The children were brought out of class to a seating area where the principal investigator was sitting with a laptop. It was explained to the children that they were going to watch a short cartoon, and then they would talk to KASPAR, who would ask them questions about the cartoon they had just seen. It was emphasised to the children that it was "not a test" and it was "a bit of fun", as we did not want the children to be stressed. Following this, the children watched the video and then went into another room where they would sit and talk to KASPAR. While the children were talking to KASPAR, the teaching assistant remained in the room but was as unresponsive as possible. The principal investigator was not in the interview room with the children. Each interview lasted between 5 and 10 minutes. The interviews with the children were not recorded because the purpose of this study was to gather feedback from the professional interviewer on their experience and views of the system after the trial.

8.2.3.4. Post-trial interview (the interviewer):

After all of the interviews were completed a post-trial interview took place with the psychologist to gather feedback on their experience with the system. The purpose of this interview was to establish what they liked about the system, what they did not like about the system, how they felt about this method of interviewing and whether they feel that this approach could be useful in a real world setting. The questions asked were as follows:

- How did you find the interviews?
- How much information did you manage to acquire from the children about the video? Was this more or less than you had expected?
- How did you feel that the children reacted to the robot?
- Was controlling the robot as you expected?
- What did you like about this method of interviewing?
- What didn't you like about this method?
- What changes would you make to the system?
- Do you think this approach could work in a real world setting?
- If you were to use this system which children do you think it would be most applicable to?
- Would a tool like this be useful in your line of work?
- Could you see any other professions where this tool could be useful?

8.2.3.5. Post-trial debrief (the children):

Two weeks after the interviews had taken place, the children were given a group debrief to show them how the robot worked and were also given the opportunity to play with and control the robot along with a number of other robotic toys. The debrief served as a thank you to the children and the school for taking part in the study, and was also educational in teaching the children a little bit about robots.

8.2.4. Measures

The data collection for this study was based on feedback from the expert interviewer. It was not necessary to record the interviews with the children, because the 3 prior studies had already shown how children responded to KASPAR in an interview setting, therefore this aspect did not require further investigation in this thesis. The pre and post-interviews with the psychologist were recorded and later transcribed for analysis. When analysing the transcriptions the key points were logged and are discussed in the conclusion of this chapter.

8.3. Results

8.3.1. Interviewer background (pre-trial interview)

The professional recruited to conduct the field trial was a senior psychologist with more than 20 years of experience working for Hertfordshire local authority, over 10 years of which a senior post had been held. The role of a senior educational psychologist is diverse with tasks such as:

- Assessing learning and emotional needs
- Developing and supporting therapeutic and behaviour management programmes
- Writing statements for children
- Investigative work for tribunals

At a senior level this is extended to training and supervising trainee educational psychologists, and standing in for team leader. The psychologist recruited was responsible for supporting students at several mainstream primary and secondary schools, which he attended on a weekly basis, conducting statutory work. At the time of this study the psychologist did not support any schools with special previsions, but did have experience of this in the past. The educational psychologist's substantial experience of working with and interviewing children made him an ideal candidate to test and provide feedback on a RMI system.

When interviewing the psychologist we asked what the most difficult aspects of the profession are. He explained that the most time consuming aspect of the role was the report writing, because a report has to be written on every child that is seen. Further to this he went on to explain that aside from the report writing, the real challenge is ensuring that the child's needs are met, whether this is in terms of behaviour problems, learning difficulties, or social emotional problems. The psychologist stated that the most difficult cases are tribunal cases where there is a conflict between what the parents want and what the authorities are offering, and that in these cases it is important to be as independent as possible when making judgements about what the child's needs actually are.

To learn more about this process we asked where the information in such cases would usually be sourced from. The psychologist explained that information is obtained from a number of sources including teachers, parents and the child, to create a complete picture of the child. The psychologist considered that the most important part of an investigation is the individual work with the child, because it is important to understand their perspective rather than relying on other people's perceptions of the child.

Depending on the case and the needs of the child, educational psychologists often work with people in other professions such as:

- Speech and language therapists
- Occupational and physiotherapists
- Paediatricians
- Social workers
- Very occasionally police officers

The educational psychologist's experience of working with professionals from a variety of disciplines is particularly beneficial for the purpose of this research, because he is likely to have a knowledgeable view of how a RMI system might be useful to other professionals.

Building on the most difficult aspects of the role, we asked what cases and children are the most difficult to deal with? The psychologist explained that it varies, but often secondary school children around the age of 13 and 14 who are showing very disturbed behaviour can be the most challenging because they can be very oppositional and defiant, which sometimes makes them difficult to connect with. Younger children are often much easier to work with when trying to change their behaviour. The psychologist stated that severe autism can sometimes present a challenge, particularly when trying to establish the triggers for disruptive behaviour. Typically developing children will often be able to say what is bothering them, whereas children with severe autism often cannot talk and sometimes will not even relate to you, making them very difficult to communicate with.

The psychologist explained another difficulty sometimes experienced is children that are very oppositional because they see the psychologist as an authority figure. Dealing with these children requires certain skills and sometimes humour is useful to make the child feel happy and comfortable. The psychologist stated that it sometimes helps to tell the child "I don't belong to the school, so you can tell me anything you want", particularly if the child is not getting on with their class teacher very well and is worried about getting reprimanded.

The psychologist explained that experience and knowing how to connect with the child plays a vital role. Children can be very challenging and difficult to communicate with because of unknown circumstances at home such as: separation of parents, a new baby is born in the family and the child is jealous, or dealing with loss or bereavement due to the death of a parent or sibling. There are a large number of possibilities that could be affecting the child. Two particular aspects that the psychologist believed that a RMI such as KASPAR could assist with are connecting with the child and not being seen as an authority figure. Because a robot such as KASPAR will be smaller than the child and would be considered closer to the child's age in terms of appearance, this may mean that the child will both view the robot as more of an equal to them, thus helping to connect with them and eliminating the potential of being viewed as an authority figure.

In an effort to understand the profession of educational psychology and how it is changing we asked the psychologist to describe two or three major trends in the profession. This question was particularly pertinent because some changes had recently been announced. The psychologist stated that up until now their services were free to schools because it was funded through central government to the local authority, this is however about to change and soon schools will have to pay for the services of educational psychologists themselves. Some core services such as statementing and statutory work will remain free to schools, but anything beyond this, the schools will have start paying for. The psychologist indicated that this change has three major implications:

1. It will change the relationship between educational psychologist's and their clients (schools), because if the service being provided to the client is free, the client does not have to worry about money, meaning that schools would most likely think twice before calling on these services in the future.
2. There is a lot of competition with various educational psychology services, and schools will not necessarily use the local educational psychology services when they are paying for the service and have a choice, therefore schools will now be in a position to pick and choose, potentially using independent psychologists.
3. New legislation associated with collaborative work is currently being implemented. Currently health service, paediatricians, speech and language therapists and social services do work together, however the work is still relatively independent between organisations. The new

legislation being implemented is designed to co-ordinate the work of education, health and social care to integrate each departments work with children, writing collaborative reports. This means children who currently have a statement of special needs, will have an education, health and care plan instead.

Aside from these changes there is also an additional change with regards to the ages of children that educational psychologists will be working with. Traditionally educational psychologists have worked with children from 0 to 19, after which point the children are transferred to adult services. Soon educational psychologists will be working with children from 0 to 25, meaning that educational psychologists will be working with young adults rather than just children. The psychologist believes that this change has been prompted by an increase in the number of children with special needs requiring support well beyond the age of 19, stating that even for typically developing children there's a lot more emphasis on preparing for transitions into adult life. Social and emotional development does not stop at 19 and there are big transitions into getting a job, leaving home and being independent. As a result educational psychologists are increasingly working with colleges and universities to help during that period of early adulthood. The psychologist stated that clinical psychologists work with people of any age, therefore there's beginning to be some overlap between the skills of educational psychologists and clinical psychologists.

To find out more about the role of an educational psychologist we asked what characteristics are most important for the role. The psychologist believes that it is important to have a positive attitude towards children and an understanding of how children develop. He explained that historically educational psychologists were also teacher trained, but this is no longer the case. Educational psychologists now train for 3 years rather than 1. Within the 3 years the trainees get some experience in education as a teaching assistant in a school. Because the training is over a much longer period this gives the trainees more time to develop the necessary skills that would have been obtained via teacher training. The psychologist also stated that good inter-personal skills are important because educational psychologists often work with a wide range of people in a variety of contexts. For example, educational psychologists work with different children from very young children to young adults and a variety of other people including parents, teachers and a range of professionals.

To ascertain a more detailed description of how technology is used in educational psychology we asked the psychologist to explain the role that technology plays on a day-to-day basis. The psychologist explained that all the staff are equipped with laptops that are secure and restricted because of data protection, and that the use of memory sticks is prohibited. In addition to this, educational psychologists often have internet enabled phones which are linked to email accounts. The psychologist stated that technology plays a crucial role in the communication and logging aspects of educational psychology.

We asked the psychologist how often he carries out interviews with children to ascertain how much experience he had of interviewing children. The psychologist stated that interviews with children were a daily task when working in schools because talking to children and gathering their views is part of every assessment. It is important for the psychologist to establish the views of the child with regards to school and anything else that the child needs help with to ensure that the child is happy and for example, they are not being bullied.

Because this research focuses on facilitating communication, we asked the psychologist what causes the most difficulties when communicating with children. The psychologist explained that children will sometimes view them as strangers and it is important to ensure that the children are comfortable with their presence. One approach to this problem is for the psychologist to be seen by the child talking to the teacher, showing the child that the he is accepted by the class teacher and is fairly familiar. Because the teacher is a trusted person, if the psychologist appears to be trusted by the trusted person, then the child will see that it's ok for them to talk to him.

We asked the psychologist what expectations he had of the robot that he would be using to interview the children. The psychologist believed that the novelty effect would be "tremendous" and that the children were likely to be very excited, although some children may be scared if they were very young. It was explained that all of the children had been briefly introduced to KASPAR as a group, so that they would be familiar with KASPAR. The psychologist believed that the children he would be working with between the ages of 7 and 9 would be comfortable with the robot, but thought that younger children might be scared of the robot. The psychologist also believed that the lack of response in the robots face could be quite inhibiting for some children, because sometimes children will be looking for positive non-verbal feedback, and if they are not getting it they may get worried and switch off or withdraw. The psychologist also believed that some of the children might want to play with the robot or even attack the robot to see what happens. Overall the psychologist thought that children between 7 and 9 years of age would probably be fascinated by the robot and because they have already seen it before and that it would be good fun for them.

The final question we asked the psychologist was what preconceptions he has in terms of the advantages and limitations of a RMI approach. The psychologist felt that it would depend on the child. For example, they thought the system would probably be useful for children with autism who find it really difficult to make eye contact and understand facial expressions, as they may be able to look at it much more easily. In addition to this, the novelty effect to children with autism might be particularly useful as some children with autism switch off to human voices, but may not with a robot. The psychologist believed that in the case of typically developing children the robot might take the place of a trusted doll. Children sometimes talk to their dolls and their teddy bears and tell them things that they wouldn't tell an adult. The psychologist did state however that in this type of situation it would be unlikely to happen immediately and that the first experience would most likely be about fun. This concluded the pre-trial interview before the study commenced later in the week where a post-trial interview would be conducted to gather feedback on the educational psychologists experience with the system.

8.3.2. System feedback (post-trial interview)

Immediately after the interviews with the ten children were completed, a post-trial interview was held with the psychologist to gather feedback on his experience, views and thoughts on the interviews he had conducted.

We started by asking the psychologist how he found the interviews? The psychologist said that the interviews were "fascinating", because all the children were different. The psychologist explained that the children all differed in terms of memory, understanding and vocabulary. Some of the children had an excellent vocabulary and it was very clear that they understood words such as "termites", in contrast, other children seemed to have forgotten large parts of the video, and didn't know the name of the insects. One child referred to them as "flies" whilst another referred to them as "mice". The psychologist found it fascinating to see how patchy short term memory was for the children, mentioning that it was quite difficult to draw information from some of them, possibly because of personality characteristics.

The psychologist also stated that some of the children were slightly weary of KASPAR, especially if the robot moved and made a big gesture. It was thought that this was possibly because some of KASPAR's movements were quite sharp and as a result may have startled and surprised some of the children. The psychologist did ask two of the children if they found me (being KASPAR) scary. Although the child denied finding KASPAR scary, the psychologist insisted that the child was startled at the time when KASPAR made the large gesture. Although the children had been given a group introduction to KASPAR, during this introduction the children could see that when a button was pressed, KASPAR would move. It is possible that because the children could not see this action and consequence in the interviews, it may have taken them by surprise. Therefore in pre-trial introductions, it may be better to give a full like for like experience of the robot so that the children would know what to expect. In addition to this it may also be beneficial to slow down some of KASPAR's movements.

The next question we asked the psychologist was, how much information did you manage to acquire from the children about the video, and was this more or less than expected? The psychologist stated that he managed to acquire much less information than expected, but believed that this may have been due to the novelty effect of the robot and nature of the activity. Although the children had been introduced to the robot prior to the study, this may not have been enough, and the novelty effect may have interfered with some of the children's memory. The psychologist believed that it would have been interesting to conduct a comparison between the robot and a teaching assistant to see if there would have been a difference in the information that the children would have provided. Although the psychologist was surprised at the lack of information he managed to elicit from the children. The previous studies with typically developing children found that there were no considerable differences in the information that the children provided. With regards to the details that the psychologist managed to acquire from the children, all of the children stated that the cartoon was a Tom and Jerry cartoon. One of the children was particularly good at recalling details of the video but still missed some major points which would have been very clear and visual in the video. The psychologist explained that even when he prompted some of the children with information he had received in interviews in an effort to get them to say more, some of the children still said that they could not remember, which surprised the educational psychologist.

In some respects this may not be negative. If the children could genuinely not remember and could not be coerced by the robot, this could be a very positive sign, indicating that children are less likely to be led by a robot. A recent study by Bethel et al. found that people were less likely to be misled by robotic interviewers than human interviewers [16], which would support this idea. The psychologist explained that as the interviews progressed he began to ask the children more questions about themselves. The psychologist stated that children are used to people they meet asking them questions such as: "do you have any brothers or sisters?", "do you have any pets?" and "do you like school?", because this is the type of language and questions that reassures the children. The psychologist believed that the children needed to be "warmed up" before going into questions about the video, and stated that the children would have benefited from a longer introduction with the robot as this may also have helped to overcome the novelty effect before the main questioning took place. The psychologist insisted that sufficient pre-interview exposure to the robot would be needed if it was to be used as an interview tool in order to get the children acclimatised to the robot so that there are no shocks.

To get a more general idea of the interaction dynamics we asked the psychologist how he felt the children reacted to the robot. The psychologist explained that there were mixed reactions, with some children who appeared to be slightly wary of the robot, whist other children really enjoyed the interaction with the robot. After the interviews one of the children mentioned to the teaching assistant that he was a little bit scared, and another child gave the impression that he was not very interested. On the contrary, some of the children seemed to enjoy the experience and were smiling and chatting all the way through the interview, although most of the children were generally somewhere between these extremes. The psychologist stated that some of the children were quite willing to talk, but their inability to remember details about the video and answer the questions seemed to put them off and were sometimes embarrassed about not being able to remember. One child in particular was interested and started talking about himself and the things in his head. The psychologist stated that this child was very happy to talk, but not just about the video. The psychologist explained that although the children were told that the activity was not a test, some children will still feel uncomfortable about not having the answer to the question they are being asked. Children of this age are particularly sensitive to this situation. The psychologist went on to say that different children have different ways of dealing with being unable to answer a question. He explained that some children will start talking more quietly, while others will start to fabricate answers, which in evidence gathering situation would obviously present a problem.

Although the study that we conducted, was with an individual who conducts interviews with children on a daily basis, the task and situation in this study was still quite artificial in terms of the children watching a short cartoon before being asked questions about it. In light of the finding that some of the children seemed to be actively willing to talk about themselves, we asked the psychologist if he felt that the children would have spoken more and in greater detail if the topic of the interview was more closely related to the child, and was something that the child was more knowledgeable about, i.e. family. The psychologist explained that children of this age are often oversensitive to not knowing something they are being asked, because they are beginning to become aware at this age that they need to start knowing things. The psychologist believed that even if the children were "warmed up" with family questions in the rapport building phase of the interview, the fact that the child is suddenly being asked a question they cannot answer will cause the problem, not the interview technique itself. The psychologist explained that younger children will often fabricate answers or not bother responding when they do not know the answer, but older children around the age of 8 and 9 will be quite sensitive about not knowing the answer to the question they are being asked.

We asked the psychologist "Was controlling the robot as you expected?". The psychologist stated that the system provided an "amazingly interesting way of remotely looking at the children from a different angle" and he quite liked this. The psychologist explained that the system took some time getting used to, but it got easier the more that he used the system, however there was a slight time delay in the system. The psychologist suggested that it would be better to be able to see KASPAR as well as the child, because it was sometimes difficult to know what position the robot was in. This would also aid in assisting the co-ordination of the robots gestures during the interaction. The psychologist stated that there was occasionally some lag time of the robots gesture and the feed from the camera that was focused on the children. The psychologist explained that eliminating the lag time coupled with having an additional camera feed focused on the robot, would make the system a lot easier to use and help to ensure the gestures were well coordinated. The psychologist also stated that because he was focused on the interviewing the child, there were a few occasions when he realised that he had not made the robot do anything for some time, by having a constant view of the robots state would help to elevate this problem. Exploring the possibilities of enhancing the system in the future, we asked the psychologist how he would feel about a system that was literally attached to the interviewer, with sensors that could detect arm and potentially head movements. The psychologist believed that this would be an excellent idea, and would make the interaction with the system from the interviewer's perspective, much more natural.

To establish the most positive aspects of the system, we asked the psychologist what he liked about this method of interviewing. The psychologist stated that he liked this method of interviewing because it was fun. The psychologist explained that he did not have any major expectations of the system and had expected more of the children not to take the robot seriously. He believed that eight out of the ten children did take the robot seriously. The psychologist stated that despite the interviews being a slightly odd, though fun situation, the children were focused on what they were being asked and actually responded as if it was an adult asking them questions, rather than going off at a tangent. The psychologist found this interesting, particularly because it very clear that the robot was a toy. The psychologist explained that the system gave him the opportunity to really focus on facial expressions of the children. When talking to someone in an interview, the interviewer cannot just stare at the person as this would make them uncomfortable. When looking via a camera from a third person perspective, it allows the interviewer to totally focus on the child's face and non-verbal feedback.

We asked the psychologist what would be the prime age group for this type of approach of interviewing. The psychologist believed that a RMI approach could be suitable for children between the ages of 5 and 7, but emphasised the importance of familiarising young children with the robot first in order to make them feel safe. However he noted that the information younger children provide may be less reliable, but it would depend on the kind of interview. The psychologist was quite adamant that a RMI approach would not be suitable for children over the age of 9.

To gather a balanced view of the system and establish which aspects of the system would need improvement, we asked the educational psychologist, "What didn't you like about this method of interviewing?". The psychologist stated that he enjoyed using the system although there were some technical difficulties with the sound and initially he couldn't hear the children clearly. The psychologist reiterated that being able to see KASPAR on the screen as well as the child would make the system much better, but there wasn't anything he particularly disliked about the system. We asked the educational psychologist, "What would you change about the system?". The psychologist stated that the sound issues would need to be resolved and that there were some lagging issues with regards to the video streaming of the child and the movements of the robot, but other than this the system was fine.

One of the most important questions we asked the psychologist was, "Do you think this approach could be used in a real world setting?". The psychologist believed that the system could be considered for real world applications provided the faults in the system were rectified first. The psychologist also confirmed that he felt the system might be useful for child protection interviews and acquiring disclosures. The psychologist explained that the system might be useful for simply getting children to talk about their worries, not necessarily for child protection issues. He stated that the children might feel able to tell the robot about things that they are worried about. The psychologist explained that some children have huge anxieties about different matters in their lives, and sometimes the children are frightened to talk about these concerns. Quite often once the child has talked about them it becomes apparent that these anxieties are about quite minor things, but once the child has expressed these concerns it is something that they can be helped with. The psychologist believed that a RMI system might be useful for highly anxious children, just to have a chat. The psychologist stated that the system would probably be suitable for children who are reluctant to talk, such as children with autism and possibly selective mutes. The psychologist explained that selective mutes, will not talk in particular settings. For example, the child may talk perfectly normally at home but will not talk at school. The psychologist explained that it is extremely difficult to get selective mutes to talk, eventually the child might whisper to a peer, or whisper to a trusted adult. The psychologist believed that it would be worth investigating the possibility of using KASPAR with that group of children, because it may be a useful tool for getting these children to talk. The psychologist stated that the system provided an interesting and totally different way of looking at children's memory. He explained that there are learning and memory tests that psychologist conduct with children, but often these tests are not very interesting for the children. By using a robot such as KASPAR this might be a good way of testing children's memory, and making the test enjoyable and interesting for the children.

The final questions we asked the psychologist were, "Would a tool like this be useful in your line of work?" and "Could you see any other professions where this tool could be useful?". The psychologist stated that he could envisage some uses for the system in his profession, explaining that finger puppets and toys are sometimes used to try and help children to talk and communicate, and that using a robot would be a step on from this. With regards to the professions where the system may be applicable, the psychologist believed that the system could potentially be useful for educational psychologists, police, social services, education, and potentially speech and language therapists. This question concluded the post-trial interview.

8.4. Conclusions

8.4.1. Summary of findings

This field trial with an educational psychologist found that similar to previous studies, most of the children responded to robot and treated the activity seriously. The psychologist was surprised by how little information many of the children could recall about the video that they had seen, however the psychologist suspects that this was not due to the robotic interviewer, rather it was the nature of the activity itself. The psychologist believed that the system would benefit from further development to resolve some of the technical difficulties such as sound and lagging issues, but could see that the system had real world potential if it was developed further. One of the potential advantages outlined by the psychologist that had not previously been considered was the ability for the interviewer to focus on the children's non-verbal feedback from a third person perspective.

When talking to someone in an interview the interviewer cannot just stare at them as this would make them uncomfortable. By a camera with a remote connection and speaking via the robot, this enables the interviewer to focus much more on the body language of the child. The psychologist believed that potential applications that this system would be useful for included child interviews with the police, social services and educational psychologists, noting that the system would probably be best suited to children with conditions that create communication difficulties such as autism. The psychologist also believed that the system may have wider applications outside of a child protection context and could be useful for children who are selective mutes and children that suffer with anxiety who find it difficult to talk to adults. An additional application area where the robot could be used outlined by the psychologist was for testing children’s memory, because using a robot would be a more interesting and engaging way to conduct these types of tests with children. Overall the feedback from the psychologist indicates that RMIs are a possibility and may have potential for real world applications. Further to this the psychologist believed that the system may also have a wider range of applications outside of child protection and encouraged us to explore these possibilities further.

8.4.2. Relation to literature and advances

The literature to date on RMI systems has primarily focused on researchers conducting studies with participants directly with their systems [13, 16]. The research conducted in this study takes the use of RMIs a step further by putting the system in the hands of a potential real world user and gathering feedback about the usability of the system and most likely applications. Gathering real world opinions from a professional who would potentially use the system in the real world is an important step in asserting if an RMI has practical applications and what would be required from the system from a usability perspective to make the transition from research to a real world system.

8.4.3. Limitations

Although this study with an educational psychologist has provided us with valuable feedback, it would have been desirable to test the system with multiple professionals from a variety of backgrounds in order to obtain views from different perspectives. In addition to this it may also have been useful to conduct multiple cycles of feedback, implementation and testing to perfect the system.

8.4.4. Review of research questions and hypotheses

The primary research question that this study aimed to answer was, RQ-7: Could robots such as KASPAR be used in an interview setting effectively by an expert working with children?. This study has shown that robots can be used by professionals to interview children, supporting the hypotheses and answering this research question. Further to this additional information has been gathered to answer the fifth and sixth research questions, "RQ-5: Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?" and "RQ-6: What would experts working with children require from a Robot-Mediated Interviewing system?". The psychologist reiterated the points stated in user panels in the previous study and also stated that the system would benefit from having a camera to monitor the status of the robot. In addition to this the psychologist outlined an additional areas where the system may be of use which were, for children that are selective mutes and for psychologist conducting memory tests because a robot may be a more interesting way to conduct the test with the children.

Chapter 9: Conclusion

9.1. Summary of research

When initially conducting a literature review surveying the current practices for interviewing children and exploring the inherent difficulties of this task for professional interviewers, we found that the two main difficulties experienced by professional interviewers were child comfort and the controllability of the interviewers own body language and facial expressions. Further to this we found that one of the most difficult groups of children to interview are children with special needs such as autism. Literature suggested the despite these children being at a greater risk of maltreatment, prosecution rates in these cases are much lower, primarily because of the difficulties experienced in communicating and gathering evidence from these children. A literature review of social robotics indicated that both neurotypical children and children with special needs appear to respond well to robots in a variety of settings, therefore we hypothesised that children may also respond well to a robot in an interview setting. We hypothesised that the children may be more comfortable with a robotic interviewer and may view it more as a peer rather than an authoritative figure. In addition to this the robots gestures and facial expressions could be precisely and explicitly controlled, eliminating unintentional expressions or body language.

Before conducting any studies we explored various methodologies and means of analysis that would be suitable for a social robotics study. To ensure that the studies conducted were of a high standard and captured all of the relevant information for analysis, prior to conducting the studies careful consideration was given to the structuring and points of measure. A standard baseline structure for the setup of the studies was devised along with a coding scheme to ensure all relevant information was captured.

The first study that we conducted explored if children would respond to a robotic interviewer and if so how these responses compared to the children talking to a human interviewer. Twenty-two children aged between 7 and 9 took part in the study. The results reveal that the children did respond to a robotic interviewer and interacted with the robot in a very similar way to how they interacted with the human interviewer. The quantitative behaviour analysis revealed that the most notable difference between the interviews with KASPAR and the human were the duration of the interviews, the eye gaze directed towards the different interviewers, and the response time of the interviewers to the children. This study builds on the findings of Bethel et al.'s work [13] which found that children are equally likely to share a 'secret', or other valuable information, with a robot as they are with a human. This study differentiates itself from Bethel et al.'s work in terms of the interview was clearly focused on gathering information from the children and the children would have been aware of this as there were no other tasks to distract them as there were in Bethel et al.'s study. In addition to this in Bethel et al.'s study the children were told a secret, and then an attempt was made by the interaction partner to obtain the secret from the child. In this study the children were not told to keep a secret, and the children were not pressurised to reveal the correct information. Rather the children were given the freedom to answer the questions as they wished with no emphasis or pressure on honest or correct responses. By giving the children this freedom we found that the children did not need to be pressured in talking to the robot in a similar way to which they did a human interviewer as they would do this naturally, which is a positive indication for the potential of RMI systems.

Following a successful first study, the second study followed-up by investigating how twenty-three children (aged between 7 and 9) respond to questions of varying difficulty from a robotic interviewer compared to a human interviewer. Similar to the first study each child participated in two interviews, one with the human experimenter and one with the humanoid robot KASPAR, however, the main questions in these interviews focused on the theme of pets and animals and had been specifically designed to vary in difficulty. Establishing how children would respond to different types of questions and questions of varying difficulty was an essential step in establishing if Robot-Mediated Interviews (RMIs) could genuinely have real world implications. To ensure that we had supplied the children with a set of questions that varied in difficulty, at the end of each interview the children were asked to rate the difficulty of the questions along with other aspects of the experience. The results from the quantitative data analysis revealed that similar to the first study, the children interacted with KASPAR in a very similar manner to how they interacted with the human interviewer, and provided both interviewers with similar information regardless of question difficulty. This study confirmed and reinforced the results of the first study by showing that children do respond to a humanoid robot in an interview setting. This study also builds on Bethel et al.'s research by indicating that question difficulty does not appear to impact on the children's responses. In Bethel et al.'s study, the children were asked to keep a secret, and it was found that the children revealed this secret to the interviewer in every instance regardless of the interviewer [13]. However, in Bethel et al.'s study the focus was on eliciting one particular piece of information from the child, with no control for question difficulty. In this study a range of questions were asked that both varied in difficulty and were designed to elicit different types and amounts of information. In police investigations, it is important to secure as much information as possible because no single piece of information can be used on its own to build a case.

Previous work in assistive robotics for children with special needs suggested that some children with special needs such as autism respond well to robots, particularly for therapeutic and educational purposes [3, 9, 71]. In light of background literature showing that children with special needs and communication difficulties being at greater risk of maltreatment, the third study investigated the possibility of RMIs with children that have various special needs. In this study we investigated how five children with special needs aged 9 to 11 responded to a robotic interviewer compared to a human interviewer. In addition to the standard measures used in the previous two studies, an additional engagement coding method was implemented. Similar to the second study, the main questions in the interviews varied in difficulty and focused on the theme of animals and pets. The results from quantitative data analysis revealed that the children interacted with the robot in a very similar manner to how they interacted with the human interviewer. However, the children with communication needs seemed to respond particularly well to the robot. Further to this the qualitative analysis, particularly from the engagement coding, suggests that the children with specific communication difficulties were more engaged with the robotic interviewer, although the number of key-points that they revealed to a robotic interviewer was generally lower. Considering the results of this study and the literature stating that children with communication difficulties require more assistance in interviews, see Section 2.1.5, this avenue of research would likely be worth pursuing in the future as there may be potential for real world impact.

Upon the compilation of three successful studies, the next step we took was to consult potential real world users for their views on the potential of RMIs. Some research on child-robot interactions in sensitive contexts has investigated the needs of the users from a variety of perspectives [54]. However, to date all studies focusing on the potential application area of RMIs have focused on the interactions between the interviewer and interviewee without an investigation of what requirements real world users would need from such a system. In order to establish if a RMI approach would work in a real world setting, it is important to establish what the experts (specialist that conduct interviews with children) would require from such a system. To establish the needs of the users we recruited 20 experts from a variety of backgrounds to participate in three separate user panels. In these user panels we gathered the participant's views of the current system and ascertained what they would require for the system to be useful to them. The user groups we worked with were specialist child protection police officers, intermediaries who specialise in working with very young children or children with communication difficulties, and a variety of other specialists. The findings from these user panels enabled us to tailor the system before trialling it with an expert that works with children.

Current literature on RMI systems has primarily focused on researchers conducting studies with participants directly with their systems [13, 16]. The final study in this research was a field trial with an independent expert user testing the RMI system. The feedback gained from the user panels furnished us with sufficient knowledge of what real world users would require from a RMI system, in order for the system to be useable to them. The core suggestions from the user panels were implemented into the system prior to conducting the trial in an effort to make the system usable for the professional testing the system. The primary implementations were, simplification of the interface, a small degree of autonomy for blinking and mouth movement whilst talking, and most importantly a direct microphone link to the robot with voice conversion to allow the interviewer to speak directly and follow any line of questioning. The direct microphone link was the most important requirement outlined in the user panels as this has the greatest impact on the flexibility of the system. Once all modifications had been made to the system, a trial was arranged with an educational psychologist to test the system with ten children. Trialling the system with a potential real world user was a critical step in establishing if this approach could have real world implications. Gathering real world opinions from a professional who would potentially use the system in the real world is an important step in asserting if an RMI has practical applications and what would be required from the system from a usability perspective to make the transition from research to a real world system.

9.2. Review of research questions

The initial purpose of the research was to establish if humanoid robots such as KASPAR could be useful tools for interviewers interviewing children. To address this overall research question there were a number of smaller research questions that needed to be answered, some of these were obvious from the outset whilst others became apparent throughout the course of the research.

RQ-1) To what degree do children respond to a robot in an interview setting?

When starting this research no prior investigations had taken place to establish if children would even respond to a robot in a formal interview setting, therefore this was the first question the research aimed to answer. The first study outlined in Chapter 4 with twenty-two children confirmed that children would respond to a robot in an interview setting. The subsequent studies in Chapters 5, 6 and 8, also provided further evidence to confirm that children do respond to a robot in a formal interview setting.

RQ-2) How do children respond to a robot compared to a human interviewer?

Establishing if children would respond to a robot in an interview setting would not be enough to ascertain if a robot could genuinely be useful in this setting. For a robot to be useful in a real world setting the children would need to respond to a robot equally or better than they would with a human interviewer, particularly in terms of information recovery. The findings of the first three studies in Chapters 4, 5 and 6 indicated that there was generally very little difference in the information that the children provided and generally the children interact with the robot in a similar way to which they would with a human interviewer.

RQ-3) Will question type and difficulty impact on children's responses to a robot in an interview setting?

Once it was established that children responded to a robot and they responded in a similar manner to which they did a human interviewer, it became necessary to ascertain how the children would respond to different types of questions and different difficulty questions. Because the nature of the foreseen application area is inherently sensitive, establishing how children respond to different types and difficulties of questions is imperative. To ascertain if RMIs could be genuinely useful, children would need to respond to a variety of types and difficulties of questions from a robotic interviewer. The results of the second study in Chapter 5 found that again children generally responded to a range of types and difficulties of questions in a similar manner with a robot to which they did with a human interviewer.

RQ-4) How do children with special needs respond to a robot compared to a human interviewer?

Before conducting research on RMIs, many studies indicated that children with special needs such as autism respond well to robots. Whilst conducting background research on investigative techniques for interviewing young children, it quickly became apparent that children with special needs or communication difficulties were at a much higher risk of maltreatment that neurotypical children. Given that research has shown children with special needs often respond well to robots and the higher likelihood of maltreatment, it was a logical step to investigate how children with special needs would respond to a robot in an interview setting. The case study with five children in Chapter 6 found that on the whole the children responded to the robot in a similar manner to which they did to a human interviewer. However, qualitative analysis and a novel coding technique found that some of the children seemed to be more engaged with the robotic interviewer than the human interviewer. This finding was not a surprise because it is in line with other research investigating how children with special needs interact with robots.

RQ-5) Do experts working with children think that a Robot-Mediated Interview approach could be useful in their field?

Once the proof of concept had been tested the next step was to establish if potential real world users considered that a RMI approach could be useful to them. This question was answered in both Chapters 7 and 8. The user panels conducted in Chapter 7 showed that the participants considered a RMI approach could be useful in some circumstances provided that the system was flexible, user-friendly and reliable. Further to this, the educational psychologist who conducted the field trial in Chapter 8 considered that a RMI system could be useful for specific groups of children in a range of application areas.

RQ-6) What would experts working with children require from a Robot-Mediated Interviewing system?

Once it had been established that children respond well to a robot in an interview setting and that potential users considered that a RMI system could be useful in their field, the next step was to ascertain what real world users would want from a RMI system. The 3 user panels outlined in Chapter 7 found that the most important features demanded by potential users were flexibility, ease of use and reliability. The participants stated that the system would need to be flexible in terms of the ability to respond to any potential line of enquiry with the children, as the nature of the interview with children is often unpredictable. In addition to this the participants also stated that the system must be user friendly so the efforts of the interviewer can focus on the interview rather than operating the system and also that the system must be reliable so the interviewer can have confidence in the system. Further to these requirements the educational psychologist that conducted the field trial in Chapter 8 gave some additional suggestions of requirements. The educational psychologist stated, that the gestures of the robot also need to be carefully considered, particularly the speed of gestures as some of them were a little fast and startled the children. In addition to this the educational psychologist stated that it would be useful to see the robot as well as the child so that the user is fully aware of the robots state.

RQ-7) Could robots such as KASPAR be used in an interview setting effectively by an expert working with children?

The final phase of this research involved implementing the requirements outlined by the user panels and testing the system with a professional user to see if it worked in practice. Chapter 8 presents the findings of a field trial with an experienced educational psychologist using the system to interview ten children about a 3 minute cartoon they watched prior to the interview commencing. In this study the psychologist successfully used the system to interview the children, illustrating that the system could be used by professional nontechnical users. For a RMI system to be useful in the real world, it would need to be sufficiently flexible and user friendly in order to be useable by non-technical users. The psychologist that conducted the study considered that the system could have real world implications in a variety of settings provided the system was developed further.

9.3. Original contributions to knowledge

Interviewing children can be a very challenging task for professionals, particularly where questions of a sensitive or emotionally provocative nature are being asked. The research outlined in this thesis aims to establish if robots such as KASPAR could be useful tools for experts working with children to conduct interviews. In conducting this research the following contributions have been made:

Human-Robot Interaction

The research conducted in this thesis has built on social robotics research by shown that children will respond to a humanoid robot in a formal interview setting. Although studies by Bethel et al. have investigated if children are willing to share secrets with robots [13], it may not have been clear to the children that then purpose of their interaction was for information acquisition as it was in the studies conducted as part of the research in this thesis. In all four of the studies conducted at schools, the children always responded to the robot. This contribution of knowledge builds on Bethel et al.'s work and is likely to be of interest to HRI researchers as very little research has investigated how robots can be utilised in an interview setting.

Further to this, we have found that children respond to a humanoid robot in a similar manner to which they would with a human interviewer regardless of question type or difficulty. In all of the studies with children the quantitative data suggested that the children responded to the robot in a similar manner as they would a human interviewer. The second study in particular investigated the impact of question difficulty and found that question difficulty does not appear to affect the interaction with a robotic interviewer in terms of information recovery. This builds on Bethel et al.'s work by establishing that question difficulty does not appear to have a large effect on the interaction with a robotic interviewer, as Bethel et al.'s study did not control for this [13]. This contribution to knowledge would likely be of interest to psychologists working with children, as it could potentially provide psychologists with a consistent interface for conducting interaction studies with children, potentially eliminating the inconsistencies of a human interviewer. However human interviews can have other advantages over robots, therefore the potential advantages of a robot are likely to depend on the application and the individual child, and would therefore need to be assessed on a case by case basis.

For robotics research to have real world impact it is important that HRI aspects are investigated from all users' perspectives. The first two HRI contributions to knowledge focus on the children, however it is also important to consider the professionals using the system. In relation to the literature, the ALIZ-E project focused on developing artificial intelligence (AI) for small social robots, and investigated the HRI aspects from a number of user's perspectives [54]. Because research into the possibility of RMI systems is relatively new, research investigating RMIs from a usability perspective is limited. The third HRI contribution to knowledge from the research in this thesis investigated the needs of the professionals who would potentially be using the system in order to establish what would be required from such a system to work in a real world situation as opposed to a research project. It was established that the most important features of a RMI system would be flexibility, stability and usability. This contribution is likely to be of interest to HRI researchers who are interested in developing RMI platforms, as this will provide them with valuable insights into what would be required for a system to be used in a real world setting by non-technical users.

Assistive robotics

Whilst reviewing literature of techniques for interviewing children, it quickly became apparent that communicating with children that have special needs can be particularly challenging. Further investigation of the literature suggests that children with special needs are more likely to be victims of maltreatment because often they are less able to communicate and as a result this makes them more vulnerable [6]. Since considerable amounts of research indicate that children with special needs often respond well to robots [3, 9] the possibility of using a RMI approach was also investigated. The contribution to knowledge for assistive robotics in this thesis comes from Study #3 conducted with children that have special needs. We found that some of the children appeared to be more engaged with the robot than the human interviewer, highlighting that this approach may have some potential. The case study conducted was with a small group of children with various special needs, therefore no specific conclusions can be drawn from this investigation. However this study coupled with the findings from previous research of assistive robotics indicates that a RMI approach may be worth further investigation and may yield some positive results. This contribution to knowledge is likely to be of interest to many groups including HRI researchers, psychologists and professionals working with children that have special needs.

Child interview communication aids

Overall this research has investigated the possibility of using a new communication aid in interviews with children. The 2011 Achieving Best Evidence document used by the UK police force, lists: drawings, pictures, photographs, symbols, dolls, figures and props, as potentially useful interview aids [7],p.89. The research in this thesis builds on this and has investigated how humanoid robots such as KASPAR could potentially be utilised as an interview aid by experts to conduct interviews with children, particularly when conventional methods are proving to be unsuccessful. A combination of the first three studies conducted, the feedback from user panels and field trial with a potential real world user all indicate that a RMI approach may have some potential, particularly in cases where children are not responding to traditional methods. This contribution to knowledge is likely to be of interest to anyone who conducts interviews with young children such as educational psychologists, social service workers and possibly police officers.

9.4. Limitations

The research in this thesis investigated how children would respond to a humanoid robot in an interview setting, with a view to establishing if robots could potentially be useful tools for professionals interviewing children. The studies in this thesis have outlined that children do respond to robots in a similar manner to which they do an adult in an interview setting regardless of the type or difficulty of the questions. Further to this the final field trial study with an educational psychologist indicated that this is an approach that may potentially be useful to professional users in a real world setting. However, there are some limitations to the research presented in this thesis, which are as follows[10]:

9.4.1. Geographical location

The children from all three of the studies were from schools in one geographical location. Children from other locations, or different cultural backgrounds, may respond differently to a robot in this setting as cultural differences may have affected the way the children would respond in the interviews.

[10] The overall limitations of this research are outlined in this section, specific limitations relating to particular studies are discussed in each chapter.

9.4.2. Age Range

The children that took part in these studies were between the ages of 7 and 11. However, the educational psychologist that conducted the final field trial with the system considered that the system might be more suitable for younger children. Therefore future work should consider investigating how such an approach would work with younger children.

9.4.3. Introversion/extroversion of children

In these studies we did not assess the children's degree of introversion or extroversion. In future research it may be useful to establish these characteristics of the child's personality and see if this affects how the children respond to a robot. Previous studies have shown the influence of participants' personality traits in human-robot interaction, [77, 110, 111].

9.4.4. Nature and duration of interviews

The questions used in the interviews and the information that the children were disclosing was not a 'personal secret' and there was no incentive for the child to keep anything from the interviewer. If the children had a vested interest in keeping information from KASPAR, or if the information had been of a more sensitive nature the results between the human interviewer and the robotic interviewer may have been different. However conducting a study that focuses on questions of a personal matter could be intrusive and would be ethically questionable. The long-term goal is therefore to develop KASPAR further as a tool for practitioners, such as members of the police or social services, rather than conducting such studies ourselves. Such future studies 'in the field' are necessary to confirm the results obtained in this research. The results from these studies provide preliminary evidence that robots could be useful tools for interviewing children.

9.4.5. Duration of interactions

This research has focused on short-term one-off interviews rather than investigating long-term child-robot interactions. This is because the initial target application area is often a novel one-off situation and children generally do not have interviews on a regular basis, therefore, long-term child-robot interactions were less relevant in the target application domain. However, future research could investigate the long-term effects for other potential interview applications, e.g. in a medical or educational context, particularly in light of the feedback we received in the user panels and from the educational psychologist suggesting that the system may be useful in a number of other areas. If robots were to be used in more long term contexts it would be important to address questions such as, will the children's behaviour differ if they are interviewed by the robot on a regular basis, and will their interest in and their co-operation with the robot decline due to the wearing off of the novelty effect?

9.4.6. Controlling of human eye gaze

The user panel of intermediaries explained that adult eye gaze can be used as a mechanism to control children and that their staff are specifically trained how to use their eyes when conducting interviews with children. In these studies the principle investigator had not been trained to use specialist eye gaze techniques. If the principle investigator had received specialist training, this may have affected how well the children responded in the first three studies in the human condition.

9.4.7. Logging agreement for key points listed in user panels

When conducting the user panels although all of the key points were logged, the number of people agreeing with each point was not recorded, therefore statistical information could not be derived from these panels.

9.4.8. Alternative means of analysis

The intermediaries and interdisciplinary user panels both commented that a qualitative analysis approach may be better suited for this research, and that perhaps a structured observation would be a better means of capturing more detailed information.

9.4.9. Professional interviewers

Although the field trial study with an educational psychologist provided us with valuable feedback, it would have been desirable to test the system with multiple professionals from a variety of backgrounds in order to obtain views from different perspectives.

9.4.10. Single inter-rater reliability checking

The data coded in each study was checked for reliability by a second video coder, and although the results of the second coder indicated that the data coded was reliable and consistent, it would have been desirable to have additional reliability coders code the videos and verify the reliability of the coding further.

9.5. Ethical considerations

This research into RMIs does have some potential ethical implications that will require consideration in the future. When conducting the interviews in the studies, the children were unaware that KASPAR was being controlled by an adult. The interviews were not of a sensitive nature and the information recovered from the children was not personal. In addition to this the children were shown how the robot worked and that they were talking to an adult after the study. In a real world investigation there are moral implications to having a child unknowingly talk to an adult via a robot and share personal or sensitive information, particularly when this information could have a significant impact on the child's life. In this research we have tested the concept of RMIs but have not fully investigated the ethical implications of using such a system in a real world setting. I would personally argue that if a using a RMI system allowed an interviewer to elicit information from a child that the interviewer would not normally have been able to obtain information from and that this information would help safe guard a child, and thus prevent this child from physical, emotional or other abuse, then this could be considered a priority. In this thesis we are investigating the potential for this approach rather than investigating ethical implications of such a system. It would ultimately be a decision for the professionals using the system to decide on the ethical use of such a system.

9.6. Future work

The results of this research have found that children respond to a robot in an interview setting and experts considered that RMIs could be useful when working with some children. With regards to future work there are a lot of potential avenues of investigation to pursue.

9.6.1. Robot-Mediated Interviews in police and social service settings

This research has tested the proof of concept, however, if robots were to be used in a police or social services type setting, further studies would need to be conducted with professionals to see how children respond to questions of a more sensitive and personal nature.

9.6.2. Robot-Mediated Interviews in a healthcare setting

During the third user panel, one of the participants indicated that using this type of system might be useful in a healthcare type environment. The example given was of children that feel like they are unable to talk about their illnesses to other people because it causes too much upset, particularly in terminal cases. The participant explained that the robots lack of emotion may be useful in this type of situation because the robot would not appear to get upset when talking to the child and could provide the child with an outlet to express themselves.

9.6.3. Robot-Mediated Interviews to assist with giving selective mutes a voice

When gathering feedback from the educational psychologist after the field trial, one of the interesting possibilities that he mentioned was using the robot to assist in getting selective mutes to talk. Selective mutes are children that can talk but cannot talk in certain settings when particular people are around them. For example, some children may talk quite happily at home but find it impossible to speak at school. The educational psychologist believed that using a robot such as KASPAR may be helpful in this type of situation because it would not be like talking to a real person and that this might gradually enable the children to speak freely at school.

9.7. Concluding remarks

The research in this thesis has investigated the possibility of Robot-Mediated communication with children in an interview setting. The foreseen goal of this research was not investigating ways of replacing human interviewers, but to provide professionals with a robotic tool that could serve as an interface to create an enjoyable and comfortable setting for children to talk to. The studies conducted during this research have illustrated that both neurotypical children and children with special needs will respond to a humanoid robot in an interview setting. Further to this when the results of these studies were presented to experts that work and conducted interviews with children, the general consensus of opinion was that a RMI system may be useful when interviewing some children, particularly children with special needs. The final field trial conducted with an educational psychologist using the system confirmed that a RMI system could be used effectively by a potential user, implying that this technology could have real world implications if further research and development was conducted.

References

[1] Aurora Project. URL: http://www.aurora-project.com/.
[2] K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, “KASPAR: A minimally expressive humanoid robot for human–robot interaction research,” *Applied Bionics and Biomechanics,* vol. 6, pp. 369-397, 2009.
[3] B. Robins, K. Dautenhahn, and P. Dickerson, "From isolation to communication: A case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot." pp. 205 – 211.
[4] J. Wainer, K. Dautenhahn, B. Robins, and F. Amirabdollahian, "Collaborating with Kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism." pp. 631-638.
[5] L. Jones, M. A. Bellis, S. Wood, K. Hughes, E. McCoy, L. Eckley, G. Bates, C. Mikton, T. Shakespeare, and A. Officer, “Prevalence and risk of violence against children with disabilities: a systematic review and meta-analysis of observational studies,” *The Lancet*, 2012.
[6] Mencap, “Submission to the Sex Offences Review,” 1999.
[7] UK Government, "Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Using Special Measures," Home Office Criminal Justice System, 2011.
[8] W. A. Bainbridge, J. W. Hart, E. S. Kim, and B. Scassellati, “The benefits of interactions with physically present robots over video-displayed agents,” *International Journal of Social Robotics,* vol. 3, no. 1, pp. 41-52, 2011.
[9] B. Robins, K. Dautenhahn, R. Boekhorst, and A. Billard, “Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills?,” *Universal Access in the Information Society,* vol. 4, no. 2, pp. 105-120, 2005.
[10] B. Huskens, R. Verschuur, J. Gillesen, R. Didden, and E. Barakova, “Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention,” *Developmental neurorehabilitation*, no. 0, pp. 1-12, 2013.
[11] V. Turk, and H. Brown, “The sexual abuse of adults with learning disabilites: results of a two year incidence survey,” *Mental Handicap Research,* vol. 6, no. 3, pp. 193-216, 1993.
[12] M. McCarthy, and D. Thompson, “A prevalence study of sexual abuse of adults with intellectual disabilities referred for sex education,” *Journal of Applied Research in Intellectual Disabilities,* vol. 10, no. 2, pp. 105-124, 1997.
[13] C. L. Bethel, M. R. Stevenson, and B. Scassellati, "Sharing a Secret: Interactions Between a Child, Robot, and Adult."
[14] NAO Robot. "Aldebaran Robotics, http://www.aldebaran-robotics.com/ last accssed: 27/12/11."
[15] K. Asmussen, *Key facts about child maltreatment*, The Institute of Psychiatry, King’s College London, April 2010.
[16] C. L. Bethel, D. K. Eakin, S. Anreddy, J. K. Stuart, and D. Carruth, "Eyewitnesses are misled by human but not robot interviewers." pp. 25-32.
[17] F. Tanaka, A. Cicourel, and J. R. Movellan, “Socialization between toddlers and robots at an early childhood education center,” *Proceedings of the National Academy of Sciences,* vol. 104, no. 46, pp. 17954-17958, 2007.
[18] UK Government., "Achieving Best Evidence in Criminal Proceedings: Guidance on Interviewing Victims and Witnesses, and Using Special Measures," Home Office Criminal Justice System, 2007, pp. 1-34.
[19] R. Spratling, S. Coke, and P. Minick, “Qualitative data collection with children,” *Applied Nursing Research*, 2010.
[20] R. Spratling, P. Minick, and M. Carmon, “The experiences of school-age children with a tracheostomy,” *Journal of Pediatric Health Care,* vol. 26, no. 2, pp. 118-125, 2012.
[21] R. E. GEISELMAN, R. P. FISHER, and H. L. HOLLAND, “EYEWITNESS MEMORY ENHANCEMENT IN THE POLICE INTERVIEW,” *On the witness stand: controversies in the courtroom,* vol. 2, pp. 113, 1987.
[22] M. Oskarsson, and C. Kriminologi, “Är förhör via chatt polisens framtida praxis för mer korrekta vittnesuppgifter?,” 2013.
[23] D. J. Gurney, K. N. Vekaria, and N. Howlett, “A Nod in the Wrong Direction: Does Non-verbal Feedback Affect Eyewitness Confidence in Interviews? 1,” *Psychiatry, Psychology and Law*, no. ahead-of-print, pp. 1-10, 2013.
[24] P. Ekman, “Darwin, deception, and facial expression,” *Annals of the New York Academy of Sciences,* vol. 1000, no. 1, pp. 205-221, 2003.
[25] R. M. Krauss, “Why do we gesture when we speak?,” *Current Directions in Psychological Science,* vol. 7, no. 2, pp. 54-60, 1998.
[26] D. McNeill, *Hand and mind: What gestures reveal about thought*: University of Chicago Press, 1992.
[27] A. L. Bradfield, and G. L. Wells, “The perceived validity of eyewitness identification testimony: A test of the five Biggers criteria,” *Law and Human Behavior,* vol. 24, no. 5, pp. 581-594, 2000.
[28] T. Shibata, "An Overview of Human Interactive Robots for Psychological Enrichment." pp. 1749-1758.
[29] A. Kerepesi, E. Kubinyi, G. K. Jonsson, M. S. Magnusson, and A. Miklosi, “Behavioral comparison of human-animal (dog) and human-robot (AIBO) interactions,” *Behavioural Processes,* vol. 73, pp. 92-99, 2006.

[30] G. F. Melson, P. H. J. Kahn, A. Beck, B. Friedman, T. Roberts, E. Garrett, and B. Gill, “Children’s behavior toward and understanding of robotic and living dogs,” *Journal of Applied Developmental Psychology,* vol. 30, pp. 92-102, 2009.
[31] B. Barlett, V. Estivill-Castro, and S. Seymon, “Dogs or robots—Why do children see them as robotic pets rather than canine machines?,” in 5th Australasian User Interface Conference (AUIC2004), Dunedin, 2004.
[32] T. Kanda, T. Hirano, D. Eaton, and H. Ishiguro, “Interactive robots as social partners and peer tutors for children: A field trial. Human-Computer Interaction,” vol. 19, pp. 61-84, 2004.
[33] F. Tanaka, and M. Ghosh, "The implementation of care-receiving robot at an English learning school for children."
[34] J. Saldien, K. Goris, and D. Lefeber, “ANTY: the Development of an Intelligent Huggable Robot for Hospitalized Children,” *CLAWAR*, 2006.
[35] K. Goris, J. Saldien, and D. Lefeber, "Probo: a testbed for human robot interaction."
[36] M. Scheef, J. Pinto, K. Rahardja, S. Snibbe, and R. Tow, “Experiences with Sparky, a social robot,” *Socially Intelligent Agents, Multiagent Systems, Artificial Societies, And Simulated Organizations,* vol. 3, pp. 173-180, 2002.
[37] M. P. Michalowski, S. Sabanovic, and P. Michel, “Roillo: Creating a Social Robot for Playrooms,” in Robot and Human Interactive Communication, The 15th IEEE International Symposium, ROMAN, 2006.
[38] M. Fior, S. Nugent, T. N. Beran, A. Ramirez-Serrano, and R. Kuzyk, “Children’s Relationships with Robots: Robot is Child’s New Friend,” *Journal of Physical Agents,* vol. 4, no. 3, 2010.
[39] S. Shahid, E. Krahmer, and M. Swerts, "Playing with iCat: investigating children's appreciation of game plays with a social robot."
[40] S. Shahid, E. Krahmer, M. Swerts, and O. Mubin, "Child-robot interaction during collaborative game play: effects of age and gender on emotion and experience."
[41] I. Leite, C. Martinho, A. Pereira, and A. Paiva, "iCat: an Affective Game Buddy Based on Anticipatory Mechanisms (Short Paper)." pp. 1229-1232.
[42] S. Shahid, E. Krahmer, and M. Swerts, “Child-robot Interaction: Playing Alone or Together,” in CHI 2011, Vancouver, BC, Canada, 2011.
[43] S. Nishio, H. Ishiguro, M. Anderson, and N. Hagita, “Representing Personal Presence with a Teleoperated Android: A Case Study with Family,” *Association for the Advancement of Artificial Intelligence*, 2007.
[44] M. King, and P. Bearman, “Diagnostic change and the increased prevalence of autism,” *International journal of epidemiology,* vol. 38, no. 5, pp. 1224-1234, 2009.
[45] M. D. Kogan, S. J. Blumberg, L. A. Schieve, C. A. Boyle, J. M. Perrin, R. M. Ghandour, G. K. Singh, B. B. Strickland, E. Trevathan, and P. C. van Dyck, “Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007,” *Pediatrics,* vol. 124, no. 5, pp. 1395-1403, 2009.
[46] R. Jordan, *Autistic spectrum disorders: an introductory handbook for practitioners*: Routledge, 2013.
[47] R. Jordan, “Autistic spectrum disorders,” *SPECIAL TEA*, pp. 110, 2004.
[48] IROMEC Project. URL: http://www.iromec.org/.
[49] H. Lehmann, I. Iacono, B. Robins, P. Marti, and K. Dautenhahn, "'Make it move': playing cause and effect games with a robot companion for children with cognitive disabilities." pp. 105-112.
[50] S. Kiesler, A. Powers, S. R. Fussell, and C. Torrey, “Anthropomorphic Interaction with a robot and robot-like agent,” *Social Cognition,* vol. 26, no. 2, pp. 169-181, 2008.
[51] C. L. Bethel, M. R. Stevenson, and B. Scassellati, "Secret-sharing: Interactions between a child, robot, and adult." pp. 2489-2494.
[52] J. Saldien, K. Goris, S. Yilmazyildiz, W. Verhelst, and D. Lefeber, “On the design of the huggable robot probo,” *Journal of Physical Agents,* vol. 2, no. 2, pp. 3-12, 2008.
[53] Aliz-E project. http://www.aliz-e.org/.
[54] I. Baroni, M. Nalin, P. Baxter, C. Pozzi, E. Oleari, A. Sanna, and T. Belpaeme, "What a robotic companion could do for a diabetic child." pp. 936-941.
[55] R. Hirose, and T. Takenaka, “Development of the humanoid robot ASIMO,” *Honda R&D Technical Review,* vol. 13, no. 1, pp. 1-6, 2001.
[56] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu, “Robovie: an interactive humanoid robot,” *Industrial robot: An international journal,* vol. 28, no. 6, pp. 498-504, 2001.
[57] M. A. Salichs, R. Barber, A. M. Khamis, M. Malfaz, J. F. Gorostiza, R. Pacheco, R. Rivas, A. Corrales, E. Delgado, and D. García, "Maggie: A robotic platform for human-robot social interaction." pp. 1-7.
[58] G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori, "The iCub humanoid robot: an open platform for research in embodied cognition." pp. 50-56.
[59] K. Kaneko, K. Harada, F. Kanehiro, G. Miyamori, and K. Akachi, "Humanoid robot HRP-3." pp. 2471-2478.
[60] K. Kaneko, F. Kanehiro, M. Morisawa, K. Miura, S. Nakaoka, and S. Kajita, "Cybernetic human HRP-4C." pp. 7-14.
[61] H. Kose-Bagci, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, "Drum-mate: A Human-Humanoid Drumming Experience."
[62] B. Robins, K. Dautenhahn, R. Te Boekhorst, and C. L. Nehaniv, "Behaviour delay and robot expressiveness in child-robot interactions: a user study on interaction kinesics." pp. 17-24.

[63] V. Tikhanoff, A. Cangelosi, and G. Metta, “Integration of Speech and Action in Humanoid Robots: iCub Simulation Experiments,” *IEEE Transactions on Autonomous Mental Development,* vol. 3, no. 1, March, 2011.
[64] K. Kaneko, F. Kanehiro, S. Kajita, H. Hirukawa, T. Kawasaki, M. Hirata, K. Akachi, and T. Isozumi, "Humanoid Robot HRP-2."
[65] S. Sabanovic, M. P. Michalowski, and R. Simmons, “Robots in the wild: Observing human-robot social interaction outside the lab,” in 9th IEEE International Workshop on Advanced Motion Control, Istanbul, 2006, pp. 596-601.
[66] H. Kose-Bagci, E. Ferrari, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, “Effects of Embodiment and Gestures on Social Interaction in Drumming Games with a Humanoid Robot,” *Advanced Robotics,* vol. 23, no. 14, pp. 1951-1996, 2009.
[67] D. Gurney, K. Pine, and R. Wiseman, “The gestural misinformation effect : skewing eyewitness testimony through gesture,” *American Journal of Psychology*, 2013.
[68] Observer XT. Noldus, http://www.noldus.com/, last access: 28/12/11.
[69] H. Kose-Bagci, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, “Drum-mate: interaction dynamics and gestures in human–humanoid drumming experiments,” *Connection Science,* vol. 22, no. 2, pp. 103-134, 2010/06/01, 2010.
[70] B. Robins, K. Dautenhahn, C. L. Nehaniv, N. A. Mirza, D. Francois, and L. Olsson, "Sustaining interaction dynamics and engagement in dyadic child-robot interaction kinesics: Lessons learnt from an exploratory study."
[71] B. Robins, E. Ferrari, and K. Dautenhahn, “Developing Scenarios for Robot Assisted Play,” in 17th IEEE International Workshop on Robot and Human Interactive Communication - RO-MAN 2008, Munich, Germany, 1-3 August, 2008, pp. 180-186.
[72] Laerdal. http://www.laerdal.com/.
[73] J. D. Gould, J. Conti, and T. Hovanyecz, “Composing letters with a simulated listening typewriter,” *Commun. ACM,* vol. 26, no. 4, pp. 295-308, 1983.
[74] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, “Wizard of Oz studies — why and how,” *Knowledge-Based Systems,* vol. 6, no. 4, pp. 258-266, 1993.
[75] K. Dautenhahn, “Socially intelligent robots: dimensions of human–robot interaction,” *Philosophical Transactions of the Royal Society B: Biological Sciences,* vol. 362, no. 1480, pp. 679-704, 2007.
[76] A. Green, H. Huttenrauch, and K. S. Eklundh, "Applying the Wizard-of-Oz framework to cooperative service discovery and configuration." pp. 575-580.
[77] M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. Te Boekhorst, and K. L. Koay, “Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion,” *Autonomous Robots,* vol. 24, no. 2, pp. 159-178, 2008.
[78] J. Goetz, S. Kiesler, and A. Powers, "Matching robot appearance and behavior to tasks to improve human-robot cooperation." pp. 55-60.
[79] C. L. Bethel, J. L. Burke, R. R. Murphy, and K. Salomon, "Psychophysiological experimental design for use in human-robot interaction studies." pp. 99-105.
[80] C. L. Bethel, and R. R. Murphy, “Review of human studies methods in HRI and recommendations,” *International Journal of Social Robotics,* vol. 2, no. 4, pp. 347-359, 2010.
[81] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, “Development and evaluation of interactive humanoid robots,” *Proceedings of the IEEE,* vol. 92, no. 11, pp. 1839-1850, 2004.
[82] J. Fast, *Body language*: Simon and Schuster, 1970.
[83] P. Ekman, and E. L. Rosenberg, *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*: Oxford University Press, 1997.
[84] K. Gwet, “Handbook of inter-rater reliability,” *Gaithersburg, MD: STATAXIS Publishing Company*, pp. 223-246, 2001.
[85] D. E. Berlyn, “The influence of complexity and novelty in figures on orienting responses,” *Journal of Experimental Psychology,* vol. 55, no. 3, pp. 289-296, 1958.
[86] J. R. Landis, and G. G. Koch, “The Measurement of Observer Agreement for Categorical Data,” *Biometrics,* vol. 33, no. 1, pp. 159-174, 1977.
[87] D. S. Boomer, “Hesitation and grammatical encoding,” *Language and speech,* vol. 8, no. 3, pp. 148-158, 1965.
[88] E. E. Shriberg, and R. J. Lickley, "Intonation of clause-internal filled pauses."
[89] S. E. Brennan, and M. F. Schober, “How listeners compensate for disfluencies in spontaneous speech,” *Journal of Memory and Language,* vol. 44, no. 2, pp. 274-296, 2001.
[90] H. H. Clark, and S. E. Brennan, “Grounding in communication,” *Perspectives on socially shared cognition,* vol. 13, no. 1991, pp. 127-149, 1991.
[91] H. H. Clark, “Managing problems in speaking,” *Speech communication,* vol. 15, no. 3-4, pp. 243-250, 1994.
[92] S. E. Brennan, and M. Williams, “The Feeling of Another’ s Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers,” *Journal of Memory and Language,* vol. 34, no. 3, pp. 383-398, 1995.
[93] V. L. Smith, and H. H. Clark, “On the course of answering questions,” *Journal of Memory and Language,* vol. 32, pp. 25-25, 1993.
[94] G. F. Mahl, *Explorations in nonverbal and vocal behavior*: Lawrence Erlbaum Associates, Inc, 1987.

[95] L. J. Wood, K. Dautenhahn, A. Rainer, B. Robins, H. Lehmann, and D. S. Syrdal, “Robot-Mediated Interviews-How Effective Is a Humanoid Robot as a Tool for Interviewing Young Children?,” *PLOS ONE,* vol. 8, no. 3, pp. e59448, 2013.

[96] L. Wing, *The autistic spectrum: A guide for parents and professionals*: Constable, 1996.

[97] R. Jordan, *Autistic spectrum disorders: an introductory handbook for practitioners*: David Fulton Publishers London, 1999.

[98] B. Robins, K. Dautenhahn, and P. Dickerson, "Embodiment and Cognitive Learning–Can a Humanoid Robot Help Children with Autism to Learn about Tactile Social Behaviour?," *Social Robotics*, pp. 66-75: Springer, 2012.

[99] I. Fenwick, and M. D. Rice, “Reliability of continuous measurement copy-testing methods,” *Journal of Advertising Research*, 1991.

[100] J. Read, S. MacFarlane, and C. Casey, "Endurability, engagement and expectations: Measuring children's fun." pp. 1-23.

[101] P. E. Shrout, and J. L. Fleiss, “Intraclass correlations: uses in assessing rater reliability,” *Psychological bulletin,* vol. 86, no. 2, pp. 420, 1979.

[102] N. Gilbert, *Researching social life*: Sage, 2008.

[103] D. E. Shapiro, “Puppet modeling technique for children undergoing stressful medical procedures: Tips for clinicians,” *International Journal of Play Therapy,* vol. 4, no. 2, pp. 31, 1995.

[104] S. Cassell, and M. H. Paul, “The role of puppet therapy on the emotional responses of children hospitalized for cardiac catheterization,” *The Journal of Pediatrics,* vol. 71, no. 2, pp. 233-239, 1967.

[105] G. Doherty-Sneddon, “Don't look now... I'm trying to think,” *The Psychologist,* vol. 17, no. 2, pp. 82-85, 2004.

[106] D. A. Poole, and M. E. Lamb, *Investigative interviews of children: A guide for helping professionals*: American Psychological Association, 1998.

[107] M. Aldridge, and J. Wood, *Interviewing Children: A Guide for Child Care and Forensic Practitioners. Wiley Series in Child Care and Protection*: ERIC, 1998.

[108] V. Braun, and V. Clarke, “Using thematic analysis in psychology,” *Qualitative research in psychology,* vol. 3, no. 2, pp. 77-101, 2006.

[109] Screaming Bee. "MorphVOX Pro," http://www.screamingbee.com/product/MorphVOX.aspx.

[110] S. Woods, K. Dautenhahn, C. Kaouri, R. te Boekhorst, K. L. Koay, and M. L. Walters, “Are robots like people?: Relationships between participant and robot personality traits in humanrobot interaction studies,” *Interaction Studies,* vol. 8, no. 2, pp. 281-305, 2007.

[111] A. Tapus, C. Ţăpuş, and M. J. Matarić, “User—robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy,” *Intelligent Service Robotics,* vol. 1, no. 2, pp. 169-183, 2008.

Appendix 1: Publications

Study #1 – Luke Jai Wood, Kerstin Dautenhahn, Austen Rainer, Ben Robins, Hagen Lehmann, and Dag Sverre Syrdal, "Robot-Mediated Interviews-How Effective Is a Humanoid Robot as a Tool for Interviewing Young Children?," PLOS ONE, vol. 8, no. 3, pp. e59448, 2013.

Study #2 – Luke Jai Wood, Kerstin Dautenhahn, Hagen Lehmann, Ben Robins, Austen Rainer, and Dag Sverre Syrdal, "Robot-Mediated Interviews: Does a robotic interviewer impact question difficulty and information recovery?," in 12th European AAATE Conference, Vilamoura, Portugal, 2013.

Study #3 – Luke Jai Wood, Kerstin Dautenhahn, Hagen Lehmann, Ben Robins, Austen Rainer, and Dag Sverre Syrdal, "Robot-Mediated Interviews: Do robots possess advantages over human interviewers when talking to children with special needs?," in International Conference on Social Robotics, Bristol, UK, 2013.

Study #4 – Luke Jai Wood, Hagen Lehmann, Kerstin Dautenhahn, Ben Robins, Austen Rainer, and Dag Sverre Syrdal, "Robot-Mediated Interviews with Children: What do potential users think?," in AISB Symposium New Frontiers in HRI, 2014.

Note: All first complete drafts were written by the author of this thesis.

Appendix 2.1: Study #1 ethics approval document

UNIVERSITY OF HERTFORDSHIRE

FACULTY OF SCIENCE, TECHNOLOGY AND CREATIVE ARTS

M E M O R A N D U M

TO	**Luke Wood**
CC	**Kerstin Dautenhahn**
FROM	**Dr Steven Adams – Chair, Faculty Ethics Committee**
DATE	**14 March 2011**

Your project entitled:

The effectiveness of humanoid robots as social mediators in an interview scenario with young children.

has been granted ethics approval and has been assigned the Protocol Number: **1011/161**

This approval is valid as follows:

From 14 March 2011

Until 30 September 2011

Please note that if it is possible that the project may continue after the end of this period, you will need to resubmit an application in time to allow the case to be considered.

Please note that your project cannot start before the approval date.

Appendix 2.2: Study #1 interview questions

		Introduction
Q1	1	Hello, my names Kasper, what's your name?
Q2	2	How old are you?
Q3	3	Tell me about yourself?
Q4	4	Do you have any brothers or sisters?
Q5	5	Tell me about them?
Q6	6	Would you like any brothers or sisters?
Q7	7	Do you have any pets?
Q8	8	Would you like any pets?
		Main Content
		Point to recover - WHAT EVENT? (Red Nose Day - Got Talent)
Q9	9	Tell me what we are going to talk about today?
Q10	0	Tell me what the event was called?
QP	~	How about the talent event?
		Point to recover - WHAT THE EVENT WAS ABOUT? (Talent Contest)
Q11	-	Describe the event to me?
		Point to recover - WHAT TOOK PLACE PRIOR TO MAIN EVENT? (Auditions)
Q12	=	Tell me what you did before the event?
Q13	q	Tell me how you got to be in the event?
Q14	w	Explain in detail your audition?
		Point to recover - WHAT HAPPENED IN THE AUDITION PROCESS? (Who was judging)
Q15	e	Tell me about the judges?
		Point to recover - WHO JUDGED THE FINAL ACTS? (Who was judging)
Q16	r	Explain what happened in the final on Friday?
Q17	t	Tell me more about them?
Q18	y	Tell me about anyone special who was there?
		Point to recover - WHO WON & WHAT DID THEY WIN? (Who won & what they recieved)
Q19	u	Tell me all about the winner?
Q20	i	Describe for me what the winner got?
		Finishing up
Q21	o	Did you enjoy the event?
Q22	p	Well It's been nice talking to you,. I'm going to meet some of your class mates now. Have fun in your lesson, and I will see you soon.

Name	Key	Generic sayings (QUESTION)
G1	a	Why?
G2	s	Why is that?
G3	f	How come?
G4	g	How many?
G5	h	Which one is your Favourite?
G6	j	Which is your favourite?
G7	k	What are they?
G8	l	What would you like?
G9	;	Why not?

Name	Key	Generic sayings (ANSWER - Specific)
C1	\	Yes
C2	z	Yes, I do
C3	x	Yes, I will
C4	c	Yes, please
C5	v	No
C6	b	No, I Don't
C7	n	No, I won't
C8	m	No, I'm not
C9	<	Yes, I'm a robot
C10	>	No, I'm a robot

Name	Key	Generic sayings (ANSWER - General)
A1	A	That's nice
A2	S	Ok
A3	D	Did you?
A4	F	Really?
A5	G	How?
A6	H	I don't know?

Name	Key	Generic sayings (ANSWER - Other)
O1	Z	Please let go of me
O2	X	Could you come back please
O3	C	Owch, that hurts
O4	V	Go on…
O5	B	What was that then
O6	N	Could you say that again?
O7	M	I'm from the University of Hertfordshire

Appendix 3.1: Study #2 ethics approval document

UNIVERSITY OF HERTFORDSHIRE

FACULTY OF SCIENCE, TECHNOLOGY AND CREATIVE ARTS

M E M O R A N D U M

TO Luke Wood

C/C Kerstin Dautenhahn

FROM Dr Simon Trainis – Chair, Faculty Ethics Committee

DATE 5 April 2012

Your Ethics application for your project entitled:

A comparative study investigating children's responses to questions of varying difficulty asked by a robot and by a human

has been granted approval and assigned the following Protocol Number:

1112/167

with the proviso that the following words are added to your parental consent form against tick box 2 so that it reads as follows:

'I also agree that any stills and/or video sequences from these trials may be used only for scientific publication, or for presentation about the project within the scientific community (videos will not be made available online or distributed), and any resulting video will be stored in a secure environment.'

This approval is valid:

From 5 April 2012

Until 3 April 2013.

If it is possible that the project may continue after the end of this period, you will need to resubmit an application in time to allow the case to be considered.

Appendix 4.1: Study #3 ethics approval document

UNIVERSITY OF HERTFORDSHIRE

FACULTY OF SCIENCE, TECHNOLOGY AND CREATIVE ARTS

M E M O R A N D U M

TO **Luke Wood**

CC **Kerstin Dautenhahn**

FROM **Dr Simon Trainis – Chair, Faculty Ethics Committee**

DATE **28 November 2012**

Your Ethics application for a class project entitled:

A comparative study investigating children's responses to questions of varying difficulty asked by a robot and by a human

has been granted approval and assigned the following Protocol Number:

1213/17

This approval is valid:

From 1 December 2012

Until 30 October 2013

If it is possible that the project may continue after the end of this period, you will need to resubmit an application in time to allow the case to be considered.

Appendix 4.2: Engagement Coding Briefing sheet

Description/Procedure
Phase 1: View all of the videos from start to finish before coding.
Phase 2: To code the video you will move an onscreen slider up and down according to how engaged that you feel that the child is with the activity and who they are talking to. The video will run at normal speed and will not be paused, coding will occur in real time. The smile above the slider will change according to the sliders position which is determined by your fingers position on the touchscreen. Once the video has completed you will be automatically prompted to save the logged data.

Definition of engagement
You will be coding engagement which is obviously very subjective but people are often very good at knowing whether someone is engaged or not. To assist with the definition there are some examples of both engagement and disengagement below:

Engagement:
Smiles
Laughing
Concentration signs (fingers in mouth, tongue out)
Excitable bouncing

Disengagement:
Frowns
Signs of boredom (ear playing, fiddling)
Shrugs

NOTE: These are just instances of what could be considered engagement and disengagement, for this task what we would really like you to do is to use your intuition and judgement and decide for yourselves how engaged you feel that the children are throughout the videos.

Appendix 5.1: Study #4 ethics approval document

UNIVERSITY OF HERTFORDSHIRE

FACULTY OF SCIENCE, TECHNOLOGY AND CREATIVE ARTS

M E M O R A N D U M

TO **Luke Wood**

CC **Kerstin Dautenhahn**

FROM **Dr Simon Trainis – Chair, Faculty Ethics Committee**

DATE **28 November 2012**

Your Ethics application for a class project entitled:

Robot-Mediated Interviews – Humanoid robot as a tool for interviewing young children (User Panel)

has been granted approval and assigned the following Protocol Number:

1213/18

This approval is valid:

From 28 November 2012

Until 25 October 2013

If it is possible that the project may continue after the end of this period, you will need to resubmit an application in time to allow the case to be considered.

Appendix 6.1: Study #5 ethics approval document

UNIVERSITY OF HERTFORDSHIRE

SCIENCE & TECHNOLOGY

M E M O R A N D U M

TO **Luke Wood**

CC **Prof. Kerstin Dautenhahn**

FROM Dr Simon Trainis, Science and Technology ECDA Chairman

DATE **24/6/13**

Protocol number: COM/PG/UH/00013

Title of study: Robot-Mediated Interviews: a field trial with experts

Your application for ethical approval has been accepted and approved by the ECDA for your school.

This approval is valid:

From: 24/6/13

To: 31/7/14

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

Appendix 6.2: Automated eye blinking code

```
package autobehaviours;

import java.awt.AWTException;
import java.awt.Robot;
import java.awt.event.KeyEvent;
import java.util.Timer;
import java.util.TimerTask;
//import java.util.logging.Level;
//import java.util.logging.Logger;
import servos.Servo;
import servos.Servos;

public class AutoBlink implements Runnable{
    public boolean active = false;
    private Timer BlinkTimer = new Timer();
    static int BlinkMain = 5000;
    static int BlinkVar = 2000;
    Servo s = Servos.getServo("EYELIDS");
    //private Robot robot;

    public void AutoBlinking(int blinkTime) {
        if(active) {
            BlinkTimer.schedule(new AutoBlink.NextBlink(), blinkTime);
            System.out.println("Blink: " + blinkTime);        /// OUTPUT VALUES
            try {
                Robot robot = new Robot();
                // Simulate a key press
                robot.keyPress(KeyEvent.VK_B);
                robot.keyRelease(KeyEvent.VK_B);
            } catch (AWTException e) {
                e.printStackTrace();
                BlinkTimer.schedule(new AutoBlink.NextBlink(), blinkTime);
            }
        }
    }

    class NextBlink extends TimerTask {
        public void run() {
            int i=(int) (BlinkMain + ((Math.random() - 0.5) * BlinkVar));
            AutoBlinking(i);
        }
    }

    public void EnableBlink() {
        System.out.println("PRESSED");
        active = true;
        AutoBlinking(BlinkMain);
    }

    public void DisableBlink() {
        System.out.println("PRESSED");
        active = false;
    }

    public void run() {

    }
}
```

Appendix 6.3: Automated mouth movement code

```
package autobehaviours;

import java.io.*;
import javax.sound.sampled.*;
import servos.Servo;
import servos.Servos;

//public class MicMouth extends JFrame{
public class MicMouth {

  boolean stopCapture = false;
  ByteArrayOutputStream byteArrayOutputStream;
  AudioFormat audioFormat;
  TargetDataLine targetDataLine;
  AudioInputStream audioInputStream;
  SourceDataLine sourceDataLine;
  Servo s = Servos.getServo("MOUTH_OPEN");

  public void EnableMicMouth(){
    try{
      audioFormat = getAudioFormat();
      DataLine.Info dataLineInfo = new DataLine.Info(TargetDataLine.class,audioFormat);
      targetDataLine = (TargetDataLine)AudioSystem.getLine(dataLineInfo);
      new CaptureThread().start();
    }catch (Exception e) {
      System.out.println(e);
      //System.exit(0);
    }//end catch*/
  }//end captureAudio method

  public void DisableMicMouth(){
      stopCapture = true;
  }

  //Setup Audio Format
  private AudioFormat getAudioFormat(){
    float sampleRate = 8000.0F;
    int sampleSizeInBits = 16;
    int channels = 1;
    boolean signed = true;
    boolean bigEndian = false;
    return new AudioFormat(sampleRate,sampleSizeInBits,channels,signed,bigEndian);
  }//end getAudioFormat

  /*
  private void setInterval(String alerthi, int i) {
        throw new UnsupportedOperationException("Not supported yet."); //To change body of generated methods, choose Tools | Templates.
    }
    */
 //Inner class to capture data from microphone
 class CaptureThread extends Thread{
  //An arbitrary-size temporary holding buffer
  byte tempBuffer[] = new byte[10000];

  int[] ad = null;
  int x = 0;
  int rate = 2000;

  public void run(){

    byteArrayOutputStream = new ByteArrayOutputStream();
    stopCapture = false;
    try{
      targetDataLine.open(audioFormat);
      targetDataLine.start();

      //Loop until stopCapture is set by another
```

```
    // thread that services the Stop button.

    while(!stopCapture){
      targetDataLine.read(tempBuffer,0,tempBuffer.length);
      int[] audioData = null;
      int nlengthInSamples = tempBuffer.length / 2;

      audioData = new int[nlengthInSamples];
      for (int i = 0; i < nlengthInSamples; i++) {
          /* First byte is LSB (low order) */
          int LSB = (int) tempBuffer[2*i];
          /* Second byte is MSB (high order) */
          int MSB = (int) tempBuffer[2*i+1];
          audioData[i] = MSB << 8 | (255 & LSB);
          x++;
          if (x>rate) {
              Thresholds(audioData[i]);////////////////////////////////////////////////////////
              x=0;
          }


      }
    }//end while
    byteArrayOutputStream.close();
    targetDataLine.stop();
    targetDataLine.close();
  }catch (Exception e) {
    System.out.println(e);
    //System.exit(0);
    //stopCapture=true;
  }//end catch
}//end run

private void Thresholds(int audioLevel){
    //System.out.println(i);/////////////////////////////////////////////////////////////

    int tooLow = 300;
    int Low = 500;
    int Mid = 700;
    int High = 800;

      if(audioLevel<tooLow) {
          System.out.println("tooLow: " + audioLevel);
          s.setPosition(600);
      } else if(audioLevel<Low) {
          System.out.println("Low: " + audioLevel);
          s.setPosition(800);
      } else if(audioLevel<Mid) {
          System.out.println("Mid: " + audioLevel);
          s.setPosition(1000);
      } else if(audioLevel<High) {
          System.out.println("High: " + audioLevel);
          s.setPosition(1200);
      } else {
          System.out.println("HHHHHH: " + audioLevel);
          s.setPosition(1400);
      }
    /*
    if (audioLevel>300&audioLevel<1300) {
        s.setPosition(audioLevel);
    } else {
        b
    }*/
  }

 }
}
```

Appendix 6.4: Automated eye movement code

```
/*
 * To change this template, choose Tools | Templates
 * and open the template in the editor.
 */
package autobehaviours;

import static autobehaviours.AutoBlink.BlinkMain;
import java.awt.AWTException;
import java.awt.Robot;
import java.awt.event.KeyEvent;
import java.util.Timer;
import java.util.TimerTask;
import servos.Servo;
import servos.Servos;

/**
 *
 * @author lwood_000
 */
public class EyeNoise implements Runnable{

    public boolean active = false;
    private Timer EyeTimer = new Timer();
    static int EyeMain = 700;
    static int EyeVar = 500;
    static int EyeVarLR = 60;
    static int EyeVarUD = 40;
    Servo LR = Servos.getServo("EYES_LR");
    Servo UD = Servos.getServo("EYES_UD");

    int EyeDefaultLR = LR.getPosition();
    int EyeDefaultUD = UD.getPosition();

    public void AutoEye(int blinkTime) {
        if(active) {
            EyeTimer.schedule(new EyeNoise.NextMove(), blinkTime);
            System.out.println("eyepos: " + LR.getPosition());        /// OUTPUT VALUES

            int i=(int) (((Math.random() - 0.5) * EyeVarLR));
            int b=(int) (((Math.random() - 0.5) * EyeVarUD));

            LR.setPosition(EyeDefaultLR+i);
            UD.setPosition(EyeDefaultUD+b);
        }
    }

    class NextMove extends TimerTask {
        public void run() {
            int i=(int) (EyeMain + ((Math.random() - 0.5) * EyeVar));
            AutoEye(i);
        }
    }

    public void Enable() {
        System.out.println("PRESSED");
        active = true;
        AutoEye(EyeMain);
    }

    public void Disable() {
        System.out.println("PRESSED");
        active = false;
    }

    public void run() {

    }
}
```

Appendix 6.5: Automated body movement code

```
/*
 * To change this template, choose Tools | Templates
 * and open the template in the editor.
 */
package autobehaviours;

import static autobehaviours.AutoBlink.BlinkMain;
import java.awt.AWTException;
import java.awt.Robot;
import java.awt.event.KeyEvent;
import java.util.Timer;
import java.util.TimerTask;
import servos.Servo;
import servos.Servos;

/**
 *
 * @author lwood_000
 */
public class BodyNoise implements Runnable{

    public boolean active = false;
    private Timer BodyTimer = new Timer();
    static int BlinkMain = 700;
    static int BlinkVar = 500;

    //Left Arm
    static int L1V = 10;
    static int L2V = 10;
    static int L3V = 10;
    static int L4V = 10;
    Servo L1 = Servos.getServo("ARM_L_1");
    Servo L2 = Servos.getServo("ARM_L_2");
    Servo L3 = Servos.getServo("ARM_L_3");
    Servo L4 = Servos.getServo("ARM_L_4");
    int L1S = L1.getPosition();
    int L2S = L2.getPosition();
    int L3S = L3.getPosition();
    int L4S = L4.getPosition();

    //Right Arm
    static int R1V = 10;
    static int R2V = 10;
    static int R3V = 10;
    static int R4V = 10;
    Servo R1 = Servos.getServo("ARM_R_1");
    Servo R2 = Servos.getServo("ARM_R_2");
    Servo R3 = Servos.getServo("ARM_R_3");
    Servo R4 = Servos.getServo("ARM_R_4");
    int R1S = R1.getPosition();
    int R2S = R2.getPosition();
    int R3S = R3.getPosition();
    int R4S = R4.getPosition();
    //private Robot robot;

    public void AutoBlinking(int blinkTime) {
        if(active) {
            BodyTimer.schedule(new BodyNoise.NextBlink(), blinkTime);
            //System.out.println("eyepos: " + s.getPosition());      /// OUTPUT VALUES

            //Left Arm
            int v1l=(int) (((Math.random() - 0.5) * L1V));
            int v2l=(int) (((Math.random() - 0.5) * L2V));
            int v3l=(int) (((Math.random() - 0.5) * L3V));
            int v4l=(int) (((Math.random() - 0.5) * L4V));
            L1.setPosition(L1S+v1l);
            L2.setPosition(L2S+v2l);
            L3.setPosition(L3S+v3l);
```

```
        L4.setPosition(L4S+v4l);

        //Right Arm
        int v1r=(int) (((Math.random() - 0.5) * R1V));
        int v2r=(int) (((Math.random() - 0.5) * R2V));
        int v3r=(int) (((Math.random() - 0.5) * R3V));
        int v4r=(int) (((Math.random() - 0.5) * R4V));
        R1.setPosition(R1S+v1r);
        R2.setPosition(R2S+v2r);
        R3.setPosition(R3S+v3r);
        R4.setPosition(R4S+v4r);

    }
}

class NextBlink extends TimerTask {
    public void run() {
        int i=(int) (BlinkMain + ((Math.random() - 0.5) * BlinkVar));
        AutoBlinking(i);
    }
}

public void Enable() {
    System.out.println("PRESSED");
    active = true;
    AutoBlinking(BlinkMain);
}

public void Disable() {
    System.out.println("PRESSED");
    active = false;
}

public void SetDefault() {
    L1S = L1.getPosition();
    L2S = L2.getPosition();
    L3S = L3.getPosition();
    L4S = L4.getPosition();

    R1S = R1.getPosition();
    R2S = R2.getPosition();
    R3S = R3.getPosition();
    R4S = R4.getPosition();
}

public void run() {

}
}
```