Let’s talk about it: Spoken conversational search with a robot for children

Thomas Beelen
University of Twente
Enschede, The Netherlands
t.h.j.beelen@utwente.nl

Khiet P. Truong
University of Twente
Enschede, The Netherlands
k.p.truong@utwente.nl

Ella Velner
University of Twente
Enschede, The Netherlands
p.c.velner@utwente.nl

Roeland Ordelman
University of Twente
Enschede, The Netherlands
r.j.f.ordelman@utwente.nl

Vanessa Evers
NTU Institute of Science and Technology for Humanity
Singapore, Singapore
vanessa.evers@ntu.edu.sg

Theo Huibers
University of Twente
Enschede, The Netherlands
t.w.c.huibers@utwente.nl

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
CHI’20,, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
In this article, we propose a robot to assist children in finding information through conversation. The proposed robot uses clarifying questions to assist children in communicating their information need. We describe the setup of our (unpublished) pilot study that investigated how children perceive robots and the information they provide. We also describe a study we are currently developing that addresses how children’s experience and search outcomes are impacted by a robot asking clarifying questions. We compare a robot using clarifying questions to a robot that replies by directly presenting information, which is how many currently available voice assistants operate. Finally, we describe our future steps. We intend to contribute to the development of children-centered information search and robot technology.

Introduction
For children, the internet is a rich source of information on many different topics. Children can search the internet in a variety of ways, for example by using a Search Engine (SE), or a Voice Assistant (VA) (e.g. Google assistant, Amazon Alexa, Apple Siri. see [17, 10]). However, research has shown that children are poorly supported by these tools in communicating their information need [6, 1, 12, 17]. We explore conversational robots to improve children’s search experience and outcomes (target audience: ages 10 to 12). A conversation with a robot is often engaging [16] and en-

Mockup example of a conversation with ambiguous query and clafifying question (Ambiguous question inspired by [22])

Child: *Where do the giants come from?*

Robot: *Do you want to know about the Giants baseball team?*

Child: *No, just giants.*

Robot: *I found a video about giants from [...] do you want to see it?*

joyable [15], avoids typing and spelling issues (see [6, 12]), and provides the opportunity to actively assist the child with clarifying their information need through conversation. We want to explore if a robot that asks *clarifying questions* can make the search process more engaging, satisfying, and effective. Furthermore, in a pilot study we investigated the trust-relationship between child and robot, and the impact on the information that is provided.

Children's developing abilities in abstract thinking [23], and limited knowledge base [19], make it challenging to come up with keywords [11]. Current VAs offer a limited style of interaction following the query-response paradigm. This means the agent responds with a single answer based on the query, without any conversational steps in between (which could be used to clarify the search intent). This query-response interaction style limits children in the following ways: 1) They cannot ask follow-up questions (which they often assume to be supported [17, 14]). 2) They have to include all relevant context into their question, which is challenging [17, 27]. 3) They do not get assisted by the system with suggestions [5], nor with clarifying questions to better understand their information need [27]. If unsatisfactory answers are retrieved, children have to reformulate their queries which can be difficult and frustrating [27]. These issues cause children to take longer on search tasks, get frustrated while searching, abandon searches, and miss out on the benefits of finding useful information [2, 7, 27, 17].

To assist children in communicating their information needs, we explore the use of *clarifying questions* which are a particularly relevant aspect of *spoken conversational search* (SCS)[25, 29]. By posing such questions, a robot can get a better understanding of the information need and thereby present more relevant results.

When introducing any new technology, unintended negative outcomes could arise. Especially children are a vulnerable group because they find it difficult to assess credibility of online information [24]. Children are known to create social bonds with robots and tend to trust them [4]. A potential risk is that children's tendency to trust, and difficulty assessing information, may cause them to assume the information from the robot to be credible, even when it's not. We investigated this risk in a pilot study. We wanted to find out how trust in a robot influenced children's endorsement of information from the robot. Furthermore, we studied how the children perceive conversational robots in an information centered task. In this experiment, 35 children played two quizzes in which a different robot provided them with possible answers. The children responded to the quiz questions by deciding whether they endorse the robot's suggestion or not. The questions were designed around trivia that children were unlikely to be familiar with, meaning they likely answer based on their perception on the robot. Children's endorsements of robot answers were recorded (a behavioral measure), as well as questionnaire answers, and semi-structured interviews. The analysis of the pilot is currently in progress. The outcomes are expected to help responsibly design the relationship between child and robot.

Spoken conversational search and clarifying questions

Automatically generating clarifying questions requires identifying ambiguous questions, choosing an aspect to clarify, and generating the text/speech [3]. Various researchers have used online product or support forums (e.g. [26, 9, 20]), or search engine logs [28], to create models for generating clarifying questions. These studies concern a general (adult) audience, and mostly consider text-based interfaces. Concerning children, it is difficult to use query logs, or data sets due to privacy issues [8], and availability. To circumvent this issue, it may be possible to generate clarification question systems based on other sources of data. For example, researchers have been able to develop *query suggestion* systems based on data sourced from children's websites [18, 21]. Similar methods may be used for generating clarification questions for children.

In addition to a robot being able to generate questions, it is important to know when the questions should be posed, and what other actions the conversational agent should take i.e. what *interaction model* it should use. Studies on SCS between two human conversation partners reveal highly complex interaction models [25]. The complexity of conversations makes it difficult to develop and evaluate interaction models. One important step is to study the balance between the benefits and costs of answering clarifying questions. For adult users it was found that clarification questions may increase satisfaction [13], but in other cases they may be overwhelming [26]. This issue is even more unclear for children, and motivates us to study how clarifying questions impact children's searches.

Proposed study

We aim to find out if children's search task performance is indeed improved when they use a robot supporting clarifying questions. We also want to know if they find it an engaging and satisfying process. We propose a study with the following research questions:

1 How does searching with a robot using clarifying questions compare to searching with a robot using the traditional query-response paradigm?

2 How are children's satisfaction and engagement impacted by clarifying questions by a robot?

3 How is children's search task performance impacted by clarifying questions by a robot?

In our proposed study, children interact with a robot to solve a set of search tasks. A robot using clarifying questions will be compared to a robot that adheres to the traditional query-response interaction, which may require reformulations to get to a satisfying answer [27]. There are three within-subject conditions. Each participant will encounter each condition twice in a randomized order, totalling six search tasks. The conditions are:

1. **Baselines** In these conditions, the robot is not conversational but uses the query-response paradigm that current commercial voice agents generally use.
   (a) **Baseline correct** The robot provides a result that answers the query correctly.
   (b) **Baseline incorrect** The robot provides a result that does not answer the query correctly (explained below). This requires the child to reformulate their query. In case the reformulation is deemed good enough, the robot may present the correct answer.
2. **SCS with clarification question** Here the robot uses SCS, particularly clarification questions, to better understand the information need before presenting a result.

Inspired by previous work [13], the search tasks are designed to contain ambiguities. These ambiguities are used to create each condition. The robot can directly answer incorrectly by interpreting the ambiguity wrongly, or the robot can directly answer correctly by adhering to the intended meaning of the ambiguous task. Furthermore, the ambiguities provide opportunity for clarifying questions. Six tasks will be developed and presented in randomized order.

The robot will be controlled by a human operator (Wizard of Oz). A script will be developed for the operator. There are two cases where the interaction cannot be fully scripted. This is when the robot presents an incorrect answer and waits for reformulations, as well as when the child answers a clarifying question. In these cases, the operator chooses from a set of scripted responses that ensure consistency between the different sessions.

Our dependent variables are satisfaction, engagement, and task performance. Building upon existing evaluation frameworks designed for children [14], we will use emoji-based questionnaires for satisfaction and engagement. Teachers' assessments of the search task outcomes will be used as a task performance measure. The questionnaire will be carried out after each of the six tasks. Additionally, the time required for each conversational turn will be recorded. After the search tasks, a short semi-structured interview will be conducted that is about children's experience with talking with the robot. Children in the target age range (10 - 12 years) will be recruited through approaching schools.

Future work

Our goal is to explore SCS with a robot for children. We studied how the robot and the provided information are perceived by children (to be published). In our planned study, we focus on the impact of clarifying questions on children's satisfaction, engagement, and task performance. Following these studies, we continue by developing an autonomous robot building upon our earlier studies as well as related work. The robot should operate in a public setting like a museum and can gather feedback and possibly data. This data can be used to develop systems for generating clarifying questions for children, as well as learn about the conversational interaction models. Ultimately, we hope to enable children access to information in a way that suits their needs.

REFERENCES

[1] Dania Bilal. Children's Search Processes in Using World Wide Web Search Engines: An Exploratory Study.. In *Proceedings of the ASIS annual meeting* (1998), Vol. 35. ERIC, 45–53.

[2] Dania Bilal and Joe Kirby. 2002. Differences and similarities in information seeking: children and adults as Web users. *Information Processing & Management* 38, 5 (2002), 649–670. DOI: http://dx.doi.org/10.1016/S0306-4573(01)00057-7

[3] Pavel Braslavski, Denis Savenkov, Eugene Agichtein, and Alina Dubatovka. What do you mean exactly? Analyzing clarification questions in CQA. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval* (2017). 345–348.

[4] Cinzia Di Dio, Federico Manzi, Giulia Peretti, Angelo Cangelosi, Paul L. Harris, Davide Massaro, and Antonella Marchetti. 2020. Shall I Trust You? From Child–Robot Interaction to Trusting Relationships. 11 (2020). DOI: http://dx.doi.org/10.3389/fpsyg.2020.00469 Publisher: Frontiers.

[5] Stefania Druga, Randi Williams, Cynthia Breazeal, and Mitchel Resnick. "Hey Google is it OK if I eat you?": Initial Explorations in Child-Agent Interaction. In *Proceedings of the 2017 Conference on Interaction Design and Children* (2017-06-27). ACM, 595–600. DOI:http://dx.doi.org/10.1145/3078072.3084330

[6] Allison Druin, Elizabeth Foss, Leshell Hatley, Evan Golub, Mona Leigh Guha, Jerry Fails, and Hilary Hutchinson. 2009. How Children Search the Internet

with Keyword Interfaces. volume (2009), 8. Issue issue. DOI:http://dx.doi.org/doi

[7] Sergio Raúl Duarte Torres. 2014. *Information retrieval for children: search behavior and solutions*. University of Twente.

[8] J.A. Fails, M.S. Pera, O. Anuyah, C. Kennington, K.L. Wright, and W. Bigirimana. Query formulation assistance for kids: What is available, when to help & what kids want. 109–120. DOI: http://dx.doi.org/10.1145/3311927.3323131

[9] Rashmi Gangadharaiah and Balakrishnan Narayanaswamy. Natural language query refinement for problem resolution from crowd-sourced semi-structured data. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing* (2013). 243–251.

[10] Radhika Garg and Subhasree Sengupta. 2020. He Is Just Like Me: A Study of the Long-Term Use of Smart Speakers by Parents and Children. 4, 1 (2020), 1–24. DOI:http://dx.doi.org/10.1145/3381002

[11] Hilary Hutchinson, Allison Druin, Benjamin B. Bederson, Kara Reuter, Anne Rose, and Ann Carlson Weeks. 2005. How do I find blue books about dogs? The errors and frustrations of young digital library users. (2005), 22–27. Publisher: Citeseer.

[12] Hanna Jochmann-Mannak, Theo Huibers, Leo Lentz, and Ted Sanders. Children searching information on the Internet: Performance on children's interfaces compared to Google. In *SIGIR* (2010), Vol. 10. 27–35.

[13] Johannes Kiesel, Arefeh Bahrami, Benno Stein, Avishek Anand, and Matthias Hagen. Toward voice query clarification. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (2018). 1257–1260. DOI: http://dx.doi.org/10.1145/3209978.3210160

[14] Monica Landoni, Davide Matteri, Emiliana Murgia, Theo Huibers, and Maria Soledad Pera. Sonny, Cerca! Evaluating the Impact of Using a Vocal Assistant to Search at School. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (2019) *(Lecture Notes in Computer Science)*, Fabio Crestani, Martin Braschler, Jacques Savoy, Andreas Rauber, Henning Müller, David E. Losada, Gundula Heinatz Bürki, Linda Cappellato, and Nicola Ferro (Eds.). Springer International Publishing, 101–113. DOI: http://dx.doi.org/10.1007/978-3-030-28577-7_6

[15] Iolanda Leite, André Pereira, Carlos Martinho, and Ana Paiva. Are emotional robots more fun to play with?. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication* (2008). IEEE, 77–82.

[16] Jamy Li. 2015. The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. 77 (2015), 23–37. DOI: http://dx.doi.org/10.1016/j.ijhcs.2015.01.001

[17] Silvia B. Lovato, Anne Marie Piper, and Ellen A. Wartella. Hey Google, Do Unicorns Exist? Conversational Agents as a Path to Answers to Children's Questions. In *Proceedings of the 18th ACM International Conference on Interaction Design and Children* (2019-06-12) *(IDC '19)*. Association for Computing Machinery, 301–313. DOI: http://dx.doi.org/10.1145/3311927.3323150

[18] Ion Madrazo Azpiazu, Nevena Dragovic, Oghenemaro Anuyah, and Maria Soledad Pera. Looking for the

Movie Seven or Sven from the Movie Frozen? A Multi-perspective Strategy for Recommending Queries for Children. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval* (2018). 92–101.

[19] Penelope A. Moore and Alison St George. 1991. Children as Information Seekers: The Cognitive Demands of Books and Library Systems. 19, 3 (1991), 161–68. Publisher: ERIC.

[20] Hassan Sajjad, Patrick Pantel, and Michael Gamon. Underspecified query refinement via natural language question generation. In *Proceedings of COLING 2012* (2012). 2341–2356.

[21] Meher T. Shaikh, Maria Soledad Pera, and Yiu-Kai Ng. Suggesting Simple and Comprehensive Queries to Elementary-Grade Children. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (2015-12). IEEE, 252–259. DOI: http://dx.doi.org/10.1109/WI-IAT.2015.193

[22] Ju-Hyun Song, Tyler Colasante, and Tina Malti. 2018. Helping yourself helps others: Linking children's emotion regulation to prosocial behavior through sympathy and trust. 18, 4 (2018), 518–527. DOI: http://dx.doi.org/10.1037/emo0000332

[23] Janet Spavold. 1990. The child as naïve user: A study of database use with young children. 32, 6 (1990), 603–625. Publisher: Elsevier.

[24] Mariya Stoilova, Rishita Nandagiri, and Sonia Livingstone. 2019. Children's understanding of personal data and privacy online–a systematic evidence mapping. (2019), 1–19. Publisher: Taylor & Francis.

[25] Johanne R. Trippas, Damiano Spina, Paul Thomas, Mark Sanderson, Hideo Joho, and Lawrence Cavedon. 2020. Towards a model for spoken conversational search. 57, 2 (2020), 102162. DOI: http://dx.doi.org/10.1016/j.ipm.2019.102162

[26] Zhenduo Wang and Qingyao Ai. Controlling the Risk of Conversational Search via Reinforcement Learning. In *Proceedings of the Web Conference 2021* (2021). 1968–1977.

[27] Svetlana Yarosh, Stryker Thompson, Kathleen Watson, Alice Chase, Ashwin Senthilkumar, Ye Yuan, and A. J. Bernheim Brush. Children asking questions: speech interface reformulations and personification preferences. In *Proceedings of the 17th ACM Conference on Interaction Design and Children* (2018-06-19) *(IDC '18)*. Association for Computing Machinery, 300–312. DOI: http://dx.doi.org/10.1145/3202185.3202207

[28] Hamed Zamani, Gord Lueck, Everest Chen, Rodolfo Quispe, Flint Luu, and Nick Craswell. Mimics: A large-scale data collection for search clarification. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (2020). 3189–3196.

[29] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W. Bruce Croft. Towards conversational search and recommendation: System ask, user respond. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (2018). 177–186.