UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

# Child-centered Interaction and Trust in Conversational AI

Ragone, G.; Bai, Z.; Good, J.; Güneysu, A.; Yadollahi, E.

DOI
10.1145/3713043.3734471

Publication date
2025

Document Version
Final published version

Published in
Proceedings: The 24th Annual ACM Interaction Design and Children Conference

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

Citation for published version (APA):
Ragone, G., Bai, Z., Good, J., Güneysu, A., & Yadollahi, E. (2025). Child-centered Interaction and Trust in Conversational AI. In *Proceedings: The 24th Annual ACM Interaction Design and Children Conference* (pp. 1235-1238). Association for Computing Machinery. https://doi.org/10.1145/3713043.3734471

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)*

Download date:13 Nov 2026

# Child-centered Interaction and Trust in Conversational AI

Grazia Ragone
grazia.ragone@uniba.it
University of Bari Aldo Moro
Department of Computer Science
Bari, Italy

Zhen Bai
zbai@cs.rochester.edu
University of Rochester
Computer Science
New York, United States

Judith Good
j.a.good@uva.nl
University of Amsterdam
Faculty of Science Informatics Institute
Amsterdam, Netherlands

Arzu Güneysu
arzu.guneysu@umu.se
Umeå University
Computing Science
Umeå, Sweden

Elmira Yadollahi
e.yadollahi@lancaster.ac.uk
Lancaster University
School of Computing and Communications
Lancaster, United Kingdom

## Abstract

As children face global challenges, creating environments that nurture hope and empower them to shape a fair, transparent future is essential. Conversational AI systems (CAIs) offer opportunities for cognitive and emotional growth, with trust built through transparent, responsive interactions. This workshop offers participants a hands-on opportunity to analyze child-CAI interactions, bringing their own use cases alongside pre-recorded examples from five countries provided by organizers. In collaboration with a diverse group of stakeholders, the focus will be on identifying the human factors that influence trust in child-AI interactions, aiming to advance guidelines for building transparent, trustworthy conversational AI systems.

## CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**; • **HCI design and evaluation methods**; • **User studies**;

**ACM Reference Format:**
Grazia Ragone, Zhen Bai, Judith Good, Arzu Güneysu, and Elmira Yadollahi. 2025. Child-centered Interaction and Trust in Conversational AI. In *Interaction Design and Children (IDC '25), June 23–26, 2025, Reykjavik, Iceland*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3713043.3734471

## 1 Background

Integrating Artificial Intelligence (AI) into children's environments presents a transformative opportunity to inspire creativity, promote learning, and cultivate trust in technology. Trust is essential for these systems to effectively support children. By critically engaging with AI, children can develop a nuanced understanding of AI capabilities and limitations, contributing to an informed and ethical future. A growing trend in AI development focuses on conversational AI (CAI) models tailored for young learners  cite xu2022. CAI systems, powered by NLP, simulate human conversation to enhance interactive learning [4]. Although not originally designed for children, these systems are increasingly being used to support co-creation, fostering a mutual partnership between children and AI to actively shape the learning process together.

These systems enhance deeper engagement, critical thinking, and problem-solving skills, helping children navigate an ever-changing world. However, their effective implementation requires more than technical proficiency; it demands an understanding of children's cognitive, emotional, and social development [14].

Conversational AI is designed to emulate human-like interactions, making human-machine communication more intuitive and engaging. It seeks to simulate empathy using methods such as natural language processing, sentiment analysis, and machine learning [2]. In the context of children's learning and play, these systems move beyond traditional tools, acting as active collaborators. In such collaborations, both the child and the AI contribute to each other's growth—AI adapts to the child's needs and interests while the child enhances the AI's ability to foster creativity and learning [5]. Cultural differences impacts children's expectations, behaviours, and trust in AI, underscoring the need for culturally sensitive design and evaluation approaches [6]. To stimulate meaningful human-AI interaction, CAI must prioritize transparency, empathy, and responsiveness [2]. Trust is not automatic–it must be earned through ethical interactions that align with children's developmental needs, ensuring AI serves as a reliable and supporting learning partner [12]. For AI to succeed, it must be technically proficient and support children's developmental needs—such as curiosity, trust, and emotional well-being [12]. Trust is essential for these systems to effectively support children; they must develop confidence in AI as a reliable learning partner, understanding its capabilities and limitations. This trust is built through transparent, ethical interactions that respect children's cognitive, emotional and social development. To nurture creativity and critical thinking, AI must prioritize transparency, empathy, and responsiveness, creating environments where children feel safe to explore, take risks, and engage in problem-solving [7, 13]. Trust in AI is not automatic: it is shaped by interactions

Permission to make digital or hard copies of all or part 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).
*IDC '25, Reykjavik, Iceland*
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1473-3/2025/06
https://doi.org/10.1145/3713043.3734471

that allow children to understand the AI's reasoning and limitations [3, 8, 10]. Children should be equipped to critically assess AI behaviour, reinforcing informed and constructive decision-making in their interactions with technology[10].

Beyond trust, defining and evaluating AI's effectiveness as a learning collaborator requires focusing on key attributes such as reliability, safety, and ethical responsibility [10]. While existing guidelines, such as Microsoft's 10 principles for responsible AI [9], emphasize privacy and data security, further research is needed to establish holistic evaluation metrics that consider cognitive and emotional development [11]. AI must respect ethical standards and promote social collaboration, ensuring that technology enhances rather than isolates or replaces human relationships [1]. In an increasingly AI-driven world, children's ability to communicate effectively and build relationships must remain a priority. The workshop will explore human factors critical to address trust in child-AI interactions, contributing to frameworks for designing child-centered CAI systems. Participants will examine cognitive, emotional, and social influences on effective interactions. Additionally, theoretical models such as the Mutual Theory of Mind [15]), which highlight the reciprocal mental state understanding in social human-AI interactions, will be discussed as a foundation for designing transparent and responsive AI systems that build trust.

## 2 Workshop Overview

This workshop explores human factors in child-CAI interactions, focusing on trust-building and child-centered AI design. Participants will analyze pre-recorded child-CAI interactions, examining emotional and cognitive aspects like adaptability and mutual understanding. Pre-workshop studies on child-AI collaboration in five countries, approved by the university's ethics committee, will provide insights into cultural and social influences.

Participants can also bring use cases for group discussions. Hands-on activities will guide strategies for exploring children's perceptions of trust, collaboration, and emotional growth in their interactions with AI systems. The workshop will emphasize AI as a collaborative partner that enhances engagement while preserving children's autonomy, addressing challenges like balancing AI guidance with independent play through scenario-based evaluations.

## 3 Main Topics of Interest

We invite submissions on children's interactions with conversational AI, focusing on trust, transparency, and explainability. Topics of interest include: 1) Mental model formation in children; 2) Tailoring AI explanations to children's cognitive models; 3) Participatory design with children; 4) Involving children in creating explainable, transparent systems; 5) Developing metrics for trust, explainability, and reliability in child-AI interactions; 6) Reviewing and improving existing human-centered AI design frameworks for children.

We welcome empirical studies, theoretical perspectives, design methodologies, position papers, and interdisciplinary approaches that contribute to the responsible integration of conversational AI into children's environments.

**Structure and Objectives** Participants will analyze five or more real-world use cases (from five countries) based on pre-recorded child-CAI interactions. These case studies will help evaluate AI's role in children's education and social experiences, with a focus on ethical and developmental implications.

Through these use cases, we aim to: 1) Examine how trust is built and measured in users; 2) Explore the AI's role in fostering collaboration and curiosity in child-AI interactions; 3) Contribute to child-centered AI design; 4) Inform future frameworks for trustworthy practices.

**Value to the Community** This workshop provides insights into the evolving dynamics of child-AI trust, a critical area as conversational AI becomes more integrated into education and social contexts. Participants will explore how these systems influence children's perceptions, learning, and social interactions while examining cross-cultural differences. The workshop aims to contribute to frameworks for safe, responsible, and effective AI use for children.

## 4 Organizers

**Grazia Ragone** is a Postdoc Researcher at the University of Bari, Italy, her work centers on developing metrics and frameworks for collaborative technologies and human-AI interaction. She holds a PhD in Human-Computer Interaction from the University of Sussex, building on a background in psychology that shapes her focus on designing and evaluating systems for individuals with diverse abilities.

**Zhen Bai** is an Assistant Professor at the University of Rochester, USA. Her research focuses on creating embodied and intelligent technologies that promote learning and social interaction for children with diverse backgrounds and abilities.

**Judith Good** is a Professor of Human-Computer Interaction at the University of Amsterdam, Netherlands. Her research interests focus on the co-design of new technologies for children with and without disabilities. She is also interested in developing new participatory methodologies for typically marginalised populations to have greater involvement in both the design and evaluation of new technologies

**Arzu Güneysu** is an Associate Professor at Umeå University in the Department of Computing Science, Sweden. She has organized several workshops in IDC, ICRR and HRI on care and education. Her research interests include topics in HCI in healthcare and education, such as Co-design for Autism Therapy.

**Elmira Yadollahi** is an Assistant Professor at Lancaster University, UK. She holds a joint PhD in robotics and computer science from EPFL in Switzerland and Instituto Superior Técnico in Portugal. Her research tackles explainability in robotics, multimodal perception and explanation to enhance the robot's transparency. She has organized a series of Workshops on Child-Robot Interaction at ICSR2020, IDC2021, IDC2024, and on Interactive AI at RO-MAN2024. She is an associate editor of the IJCCI Journal and has served as chair and AC in IDC, ICSR, AAMAS, and HRI.

## 5 Website

A dedicated website (http://c3ai.di.uniba.it) has been published, providing all the relevant information about the workshop, including the call for participation, key dates, format, submission instructions, and organizer biographies. Once the selection process for submitted papers is finalized, the website will be updated to include a list of

position papers, a conference programme, a list of participants and a link to each participant's position paper.

## 6 Pre-Workshop Plans

In the lead-up to the workshop, we will video record at least five child-AI conversations (ages 7-11) from five countries across three scenarios: 1) Educational (e.g., math), 2) Playful (e.g., games or storytelling), and 3) Problem-solving (e.g., puzzles). These interactions will help analyze how trust is built and understood by children in different contexts.

The workshop will include a hands-on analysis of these interactions, allowing participants to explore factors influencing children's trust, engagement, and emotional responses to AI. To maximize participation, we will use a dedicated website, active social media, and collaborate with educational institutions to reach AI, HCI, and child development professionals. Engagement in relevant online forums will further spread awareness. We will encourage content contributions, fostering a collaborative atmosphere. A call for position papers will invite diverse perspectives. Through these initiatives, we aim to cultivate a vibrant and interconnected community prior to the workshop.

## 7 Workshop Structure

The workshop spans half a day, with the following schedule:

**09:00 - 09:15** Welcome and introduction to the workshop
**09:15 - 9:30** Speed networking
**9:30 - 10:30** Hands-on Activity: participants will work in small groups to analyze real child-AI interactions, evaluate key human factors, and develop strategies for enhancing trust, engagement, and transparency in AI systems for children.
**10:30 - 10:45** Coffee break
**10:45 - 12:15** Evaluation Metrics brainstorming and Panel discussion
**12:15 - 12:45** Invited Keynote : Dr Marina Escobar Planas will explore "Trustworthy Conversational Agents for Children" through the lens of the European Commission's definition of ethical AI. Her work focuses on helping children understand the limitations of these technologies, aiming to manage expectations—sometimes by intentionally challenging over-trusting perceptions.
**12:45 - 13:00** Concluding remarks

**Resources needed** We will require a room accommodating approximately 30 people with projection facilities. The room should also be suitable for dividing people into groups to discuss.

## 8 Post-Workshop Plans

Building on the workshop outcomes, we will develop a checklist for Human-Centered AI (HCAI) for children, focusing on ethical tools, interactions, and agents. Key areas include playful interactions, child-appropriate models, ethics, data security, and children's rights. Participants may also expand their contributions for potential inclusion in a special International Journal of Child-Computer Interaction (IJCCI) special issue, subject to workshop acceptance.

## 9 Call for Participation

The C3AI workshop at IDC25 invites participants to investigate the human factors involved in building trust in children toward AI, while also exploring how to encourage critical thinking and awareness of AI's limitations. We seek submissions on topics such as:

- What fosters trust towards AI;
- Designing AI systems that encourage critical thinking and awareness of AI's limitations;
- Human-centered design for children;
- Ethical considerations for AI systems;
- Assessing AI's impact on children's development.

a) what fosters trust towards AI; b) Designing AI systems that encourage critical thinking and awareness of AI's limitations; c) Human-centered design for children; d) Ethical considerations for AI systems; e) Assessing AI's impact on children's development. Submissions (1-4 pages) should present insights or research on trust-building in Child-AI interactions. Formats include: a) Abstracts (300-500 words) b) Design examples (600-800 words) c) Short papers (4+ pages) We also welcome videos and graphics. Submissions will be managed via EasyChair, and up to 20 participants will be selected based on quality and diversity. Accepted papers will be featured on the workshop website, and at least one author must attend and register for both the workshop and the main conference.

We look forward to your contributions!

## Acknowledgments

This work is partially supported by the co-funding of the European Union—Next Generation EU: NRRP Initiative, Mission 4, Component 2, Investment 1.3 – Partnerships extended to universities, research centres, companies and research D.D. MUR n. 341, 15.03.2022 – Next Generation EU (PE0000013 – "Future Artificial Intelligence Research – FAIR"—CUP: H97G22000210007)

## References

[1] Ryoko Asai. 2020. AI and ethics for children: How AI can contribute to children's wellbeing and mitigate ethical concerns in child development. In *Societal Challenges in the Smart Society*. Universidad de La Rioja, 459–466.
[2] Raymond R Bond, Felix Engel, Michael Fuchs, Matthias Hemmje, PM McKevitt, Michael McTear, Maurice Mulvenna, Paul Walsh, and Huiru Zheng. 2019. Digital empathy secures Frankenstein's monster. In *Collaborative european research conference*. Collaborative European Research Conference, 335–349.
[3] X. Chen and I. Zainuddin. 2023. Comprehensive Study of Emotional Responses in AI-Enhanced Interactive Installation. *Sustainability* 15, 22 (2023), 1–22. doi:10.3390/su152215830
[4] Google Cloud. 2025. Conversational AI. https://cloud.google.com/conversational-ai Accessed: 2025-01-21.
[5] S. Druga and R. Williams. 2017. *Children are growing up with technology that blurs the line between animate and inanimate objects. How does this interaction affect kids' development?* Technical Report 1. MIT Media Lab. https://www.media.mit.edu/posts/kids-ai-devices/
[6] Jarosław Kozak and Stanisław Fel. 2024. How sociodemographic factors relate to trust in artificial intelligence among students in Poland and the United Kingdom. *Scientific Reports* 14, 1 (2024), 28776.
[7] Nomisha Kurian. 2024. 'No, Alexa, no!': designing child-safe AI and protecting children from the risks of the 'empathy gap'in large language models. *Learning, Media and Technology* (2024), 1–14.
[8] Wenwei Luo, Huihua He, Minqi Gao, and Hui Li. 2024. Safety, Identity, Attitude, Cognition, and Capability: The 'SIACC'Framework of Early Childhood AI Literacy. *Education Sciences* 14, 8 (2024), 871.
[9] Microsoft. 2020. Responsible Bots: 10 Guidelines for Developers of Conversational AI. https://www.microsoft.com/en-us/research/publication/responsible-bots/ Accessed: 2025-01-15.

[10] Athanasios Polyportis and Nikolaos Pahos. 2024. Understanding students' adoption of the ChatGPT chatbot in higher education: the role of anthropomorphism, trust, design novelty and institutional policy. *Behaviour & Information Technology* (2024), 1–22.

[11] Grazia Ragone, Safinah Arshad Ali, Andrea Esposito, Judith Good, Kate Howland, and Carmelo Presicce. 2024. Child-Centered AI for Empowering Creative and Inclusive Learning Experiences. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. 1035–1037.

[12] Surbhi Seema Sethi and Kanishk Jain. 2024. AI technologies for social emotional learning: recent research and future directions. *Journal of Research in Innovative Teaching & Learning* (2024).

[13] S. Vosloo. 2023. *Generative AI: Risks and Opportunities for Children* . Technical Report 1. UNICEF Innocenti – Global Office of Research and Foresight. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.unicef.org/globalinsight/media/3061/file

[14] Ge Wang, Kaiwen Sun, Ayça Atabey, Kruakae Pothong, Grace C. Lin, Jun Zhao, and Jason Yip. 2023. Child-Centred AI Design: Definition, Operation, and Considerations. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 338, 6 pages. doi:10.1145/3544549.3573821

[15] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. 2021. Towards mutual theory of mind in human-ai interaction: How language reflects what students perceive about a virtual teaching assistant. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.