

Evaluating the Believability of a Conversational Agent

Author: Vladimir Makarov
v.makarov@student.tudelft.nl
Supervisors: Willem-Paul Brinkman,
Mohammed Al Owayyed

1

Introduction

- Chatbots can be used in Communication Skills Training
- Chatbots as a powerful tool for conveying personalized information
- Simulated environment for trainees to interact with chatbots



5

Research question:

How believable do users find the interaction with the Lilobot conversational agent?

- Evaluating the believability of Lilobot through a user study
- Identifying factors that may impact believability in this context

2

Subject of study: Lilobot

- Chatbot acts as a child who was bullied at school
- Previous work on usability and usefulness [1] but lacking believability measurement



6

Methodology



Viewing example conversation transcripts



Having a conversation with the chatbot



Survey based on the ASA questionnaire [2]



Interview for qualitative feedback

3

Importance of Believability

- The chatbot is acting as a real child
- Trainee's belief in the authenticity of the conversation is important
- Reducing risks and avoiding harm to human participants



7

Results

- Artificial-Social-Agent Questionnaire results suggest improvement to believability is necessary
- The mean value for the Human-Like Behaviour construct is -0.6 (SD = 1.65), and the mean for Natural Behaviour is -0.87 (SD = 1.66)
- Interview results show that use of emoticons and acknowledging the context of the application raise believability
- Unresponsiveness and repeated utterances lower believability
- Lilobot's emotions considered real and valid, especially when contextualized
- Participants suggested improving the appropriateness of its reactions and expanding its vocabulary

4

Existing Work: Believability Assessment

- Recognizing the importance of assessing the believability of conversational agents
- Highlighting the specific relevance for the agent simulating a real-life child

8

Limitations

- Chatbot is hard to work with (slow and unreliable)
- Lack of potential study participants (best would be helpline agents)
- Lack of time to teach participants good conversation techniques

[1] Sharon Grundmann. A bdi-based virtual agent for training child helpline counsellors. 2022. URL <https://repository.tudelft.nl/islandora/object/uuid%3A104f8f0b-9ab9-4f1c-a19c-43b164d45cce>.

[2] S. Fitriani, M. Bruijnes, Deborah Richards, Andrea Bönsch, and W. P. Brinkman. The 19 Unifying Questionnaire Constructs of Artificial Social Agents: An IVA Community Analysis. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents, IVA 2020*, 2020. doi: 10.1145/3383652.3423873. URL <https://repository.tudelft.nl/islandora/object/uuid%3A521411a1-579e-4054-878d-31ba3815b628>. Publisher: Association for Computing Machinery (ACM).

