Robots with Attitudes: Influence of LLM-Driven Robot Personalities on Motivation and Performance

Dennis Becker, Kyra Ahrens, Connor Gäde, Erik Strahl, Stefan Wermter

Knowledge Technology Group, Dept. Informatics, University of Hamburg, Germany

KNOWLEDGE TECHNOLOGY

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Motivation

- A robot's agreeable personality can increase cooperation and performance[1,2]
- Large language model Vicuna is used to model an **agreeable and non-agreeable personality** for a human-robot cooperative task[3,4]

- **Research questions:**

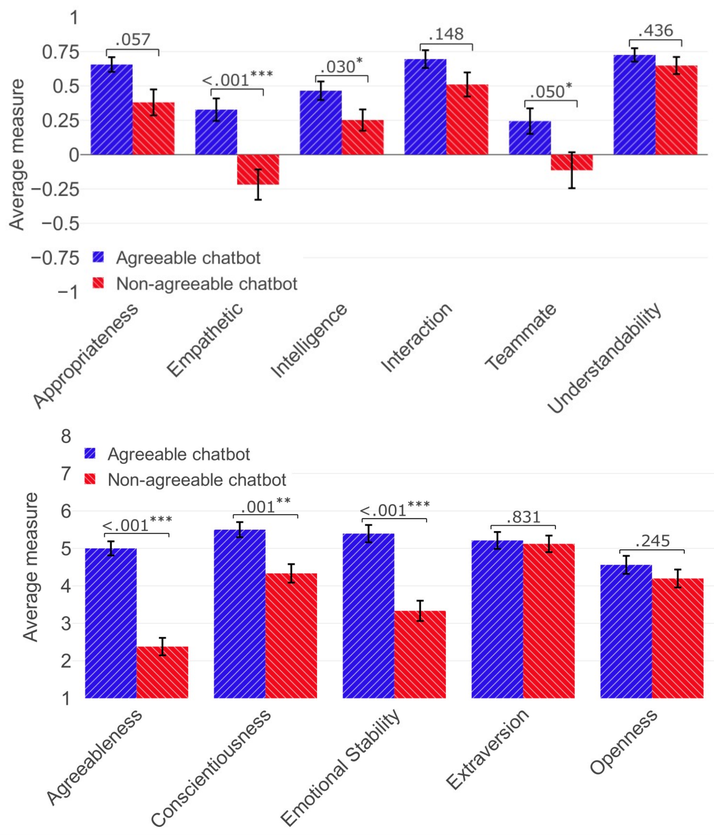
  RQ1: Can an LLM consistently convey a robot's personality?

  RQ2: Is a robot with an agreeable personality perceived as more likable?

  RQ3: Does a robot's agreeable personality increase intrinsic motivation and task performance?

Pre-study

- Perception of the chatbot
  - The agreeable personality is perceived as: **Empathetic, intelligent, and preferred as a teammate**
  - The non-agreeable personality is perceived as: **Unfriendly, less empathetic, and impedes the task**

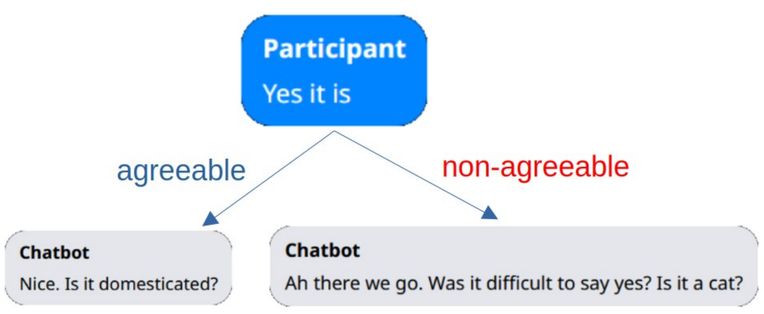


- Personality of the chatbot
  - The agreeable personality is perceived as: **Agreeable, conscientious, and emotionally stable**

Experiment design

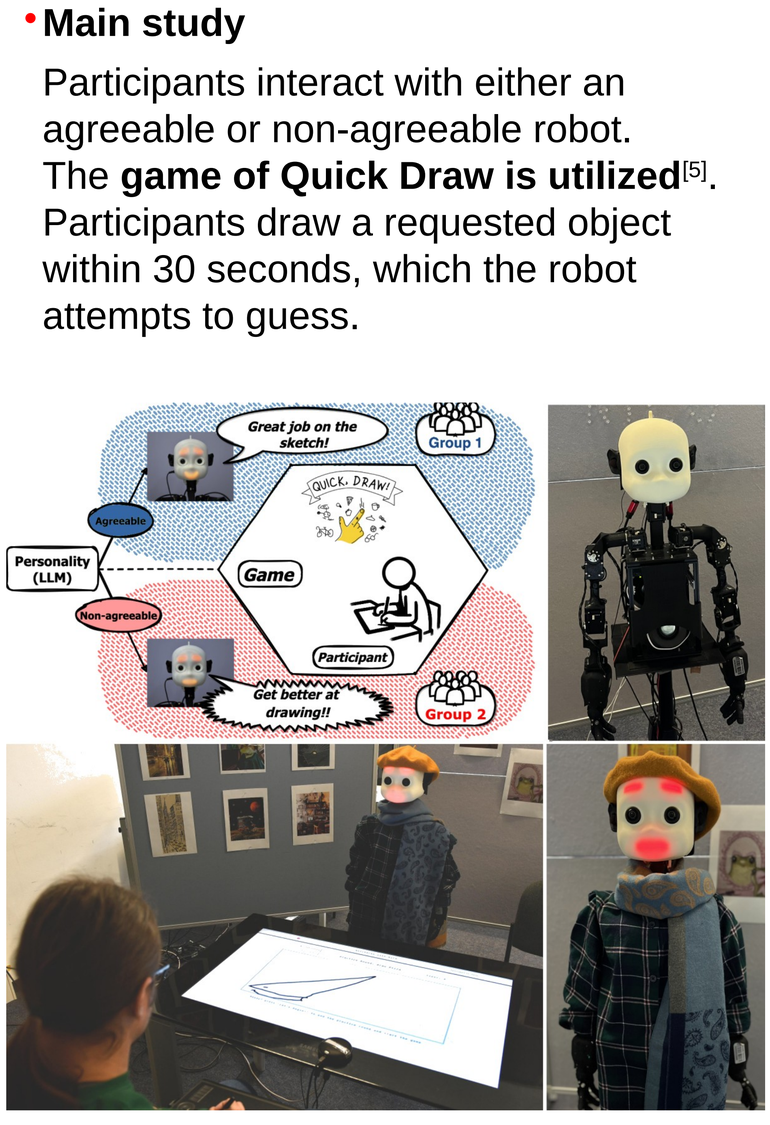
- For evaluating the influence of a robot's LLM-based personality, the study consists of two parts
  - **Pre-study**

    Participants are **presented with a chat dialog.** The chatbot needs to ask questions and correctly guess an animal.



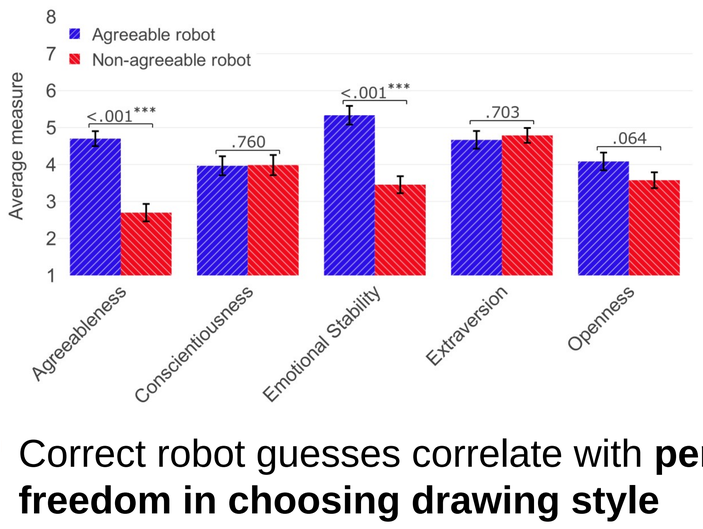
  - **Main study**

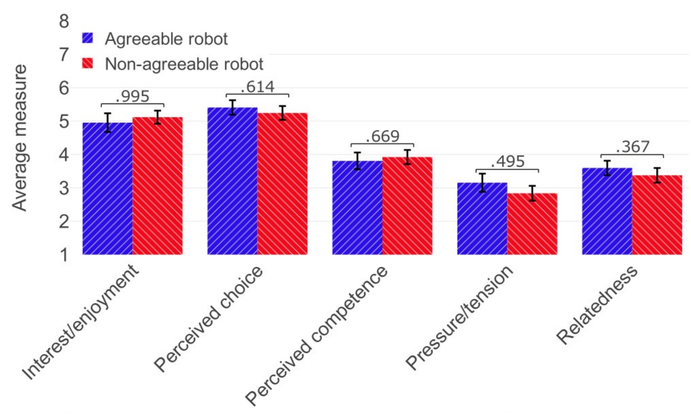
    Participants interact with either an agreeable or non-agreeable robot. The **game of Quick Draw is utilized**[5]. Participants draw a requested object within 30 seconds, which the robot attempts to guess.



Main study

- The non-agreeable robot is perceived as **less emotionally stable**, possibly due to the robot's rude comments during the interaction. **No difference in perceived motivation**





- Correct robot guesses correlate with **perceived drawing competence** and the perceived **freedom in choosing drawing style**

	Intrinsic Motivation Inventory (IMI)				
Item	Interest	Competence	Choice	Pressure	Relatedness
τ	.147	.201	.216	-.116	.146
p value	.111	.030*	.021*	.211	.115

- Perceived **robot agreeableness** and **robot openness can improve performance** for some participants

	Ten-Item Personality Inventory (TIPI)				
Item	Agreeableness	Extraversion	Conscientiousness	Emotional Stability	Openness
τ	.248	.068	.038	.037	.199
p value	.009**	.475	.689	.699	.037*

Conclusion

- **Large language models can reliably model a personality** that is recognized in an online and lab experiment (**RQ1**)
- Agreeable robots are **more likeable and preferred as teammate** (**RQ2**)
- Although a robot's agreeableness could affect task performance, a robot's personality depends on **individual preferences and the specific task** (**RQ3**)

References

Literature:

- [1] George A. Neuman and Julie Wright. 1999. Team effectiveness: Beyond skills and cognitive ability. Journal of Applied Psychology 84, 3 (1999), 376–389. doi:10.1037/0021-9010.84.3.376
- [2] Miranda A.G. Peeters, Harrie F.J.M. van Tuijl, Christel G. Rutte, and Isabelle Reymen. 2006. Personality and team performance: A meta analysis. European journal of personality 20, 5 (2006), 377–396. doi:10.1002/per.588
- [3] Antonio Andriella, Henrique Siqueira, Di Fu, Sven Magg, Pablo Barros, Stefan Wermter, Carme Torras, and Guillem Alenyà. 2021. Do I have a personality? Endowing care robots with context-dependent personality traits. International Journal of Social Robotics 13, 8 (2021), 2081–2102. doi:10.1007/s12369-020-00690-5
- [4] Wei-Lin Chiang, Zhuohan Li, Zi Lin, Ying Sheng, Zhanghao Wu, Hao Zhang, Lianmin Zheng, Siyuan Zhuang, Yonghao Zhuang, Joseph E. Gonzalez, Ion Stoica, and Eric P. Xing. 2023. Vicuna: An open-source chatbot impressing GPT-4 with 90%* ChatGPT quality.
- [5] Chipp Jansen and Elizabeth Sklar. 2021. Exploring co-creative drawing workflows. Frontiers in Robotics and AI 8, May (2021), 1–26. doi:10.3389/frobt.2021.577770

Acknowledgements

The authors gratefully acknowledge support from the German Research Foundation DFG (CML, LeCAREbot), the European Commission (TRAIL, TERAIS), and the Federal Ministry for Economic Affairs and Climate Action (BMWK) under the Federal Aviation Research Programme (LuFO), Projekt VeriKAS.