Acceptability and Trustworthiness of Virtual Agents by Effects of Theory of Mind and Social Skills Training

Hiroki Tanaka and Takeshi Saga and Kota Iwauchi and Satoshi Nakamura,

Division of Information Science, Nara Institute of Science and Technology, Japan

Socially Interactive Human-like Virtual Agents, FG2023

Social skills training (SST)

SST: evidence-based psychosocial treatment

We automated the SST process by virtual agents [Tanaka+2021; Tanaka+2022]

Collected SST data

SST for Children [Iwasaka+2010]

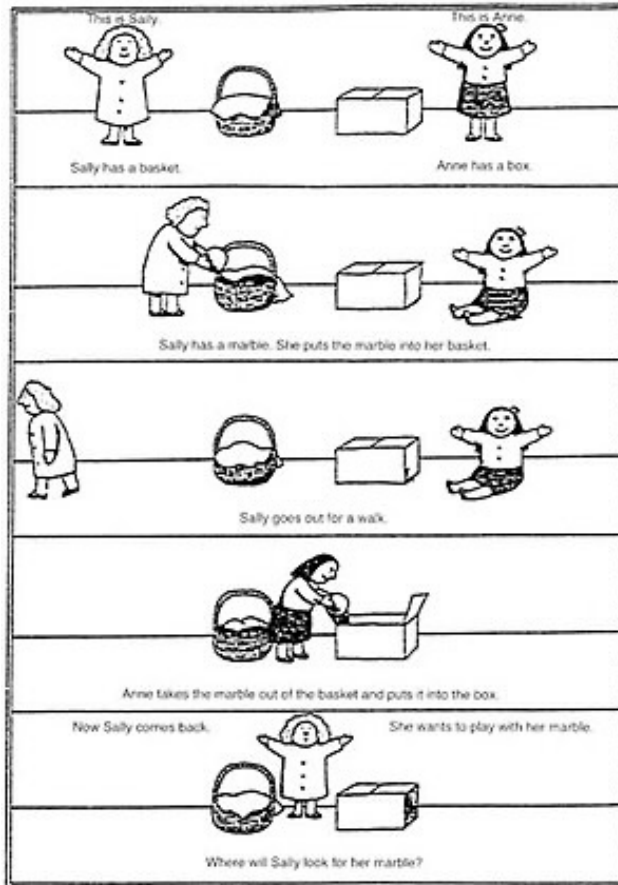# Automating social skills training [Saga+2022]

Purpose of this study

1. Whether the user's acceptability changed significantly depending on the contents: level of the virtual agent's theory of mind

2. Whether the user's acceptability improved after SST

Theory of Mind (ToM)

ToM：inference of the mental states of others, including their feelings, thoughts, and beliefs
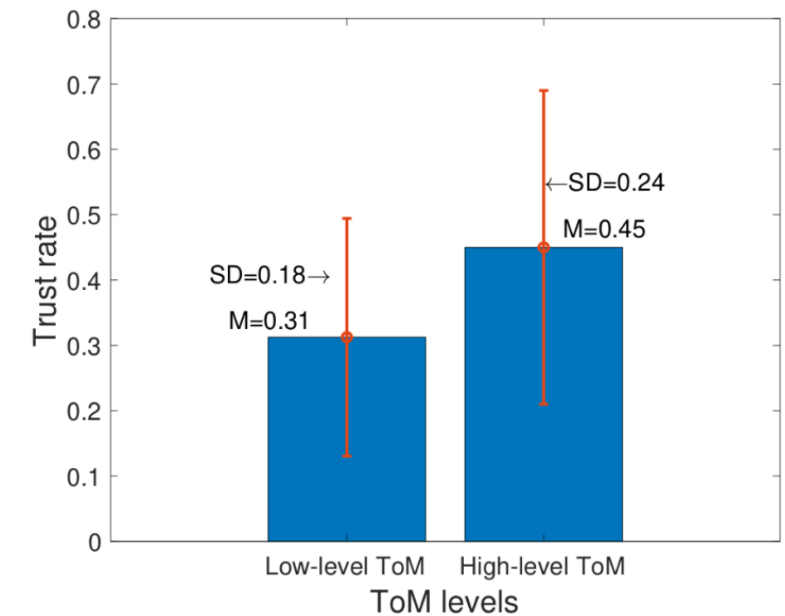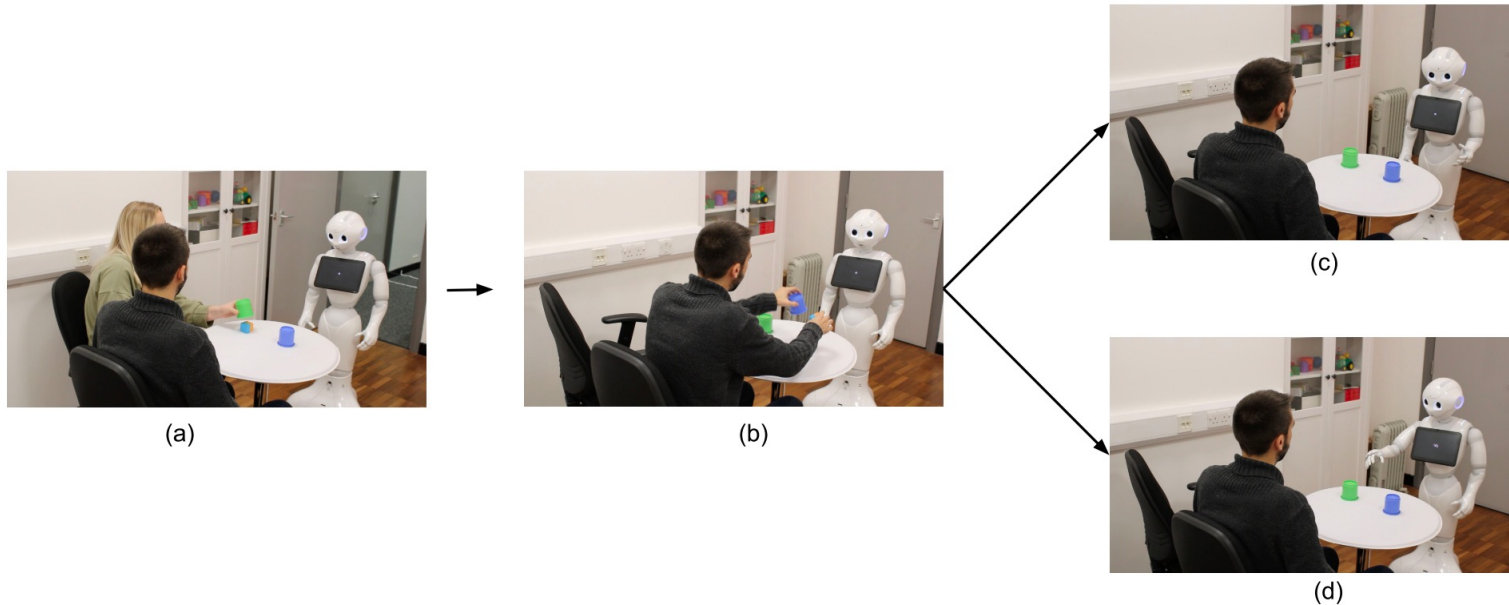

Sally and Anne test

Image: https://upload.wikimedia.org/wikipedia/en/thumb/a/ac/Sally-Anne_test.jpg/300px-Sally-Anne_test.jpg

ToM and agents

User's trust was reduced by watching a ToM video where the agent fails the modified Sally and Anne test [Mou+2021]

Clarify the following two hypotheses using our SST system:

(H1) The virtual agents' ToM levels have effects on their acceptability and trust as trainers

(H2) Distrust is weakened by the SST process

Experiment

The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of the Nara Institute of Science and Technology

Collected data from 29 participants from 22 to 35 years old (12 males and 17 females): all were Japanese without any history of psychiatric disorders

Conducted a three-person dialogue (an agent trainer and two users) and recorded videos of cases where the virtual agents failed and passed the false-belief task

Questions: acceptability as a trainer, trustworthiness, likeability, and familiarity

The participants rated each question on a 5-point Likert scale (1: completely disagree, 5: strongly agree)
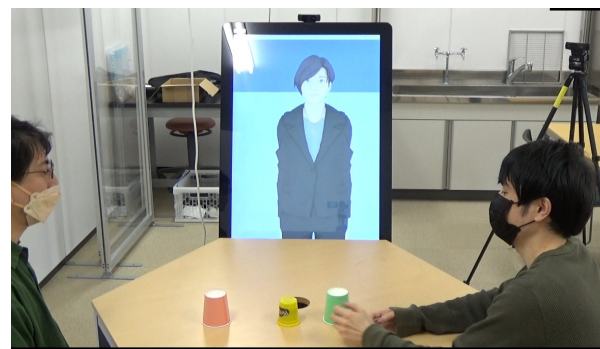
Low-level ToM (failed case)

Experiment procedure

Questions

Questions

Questions
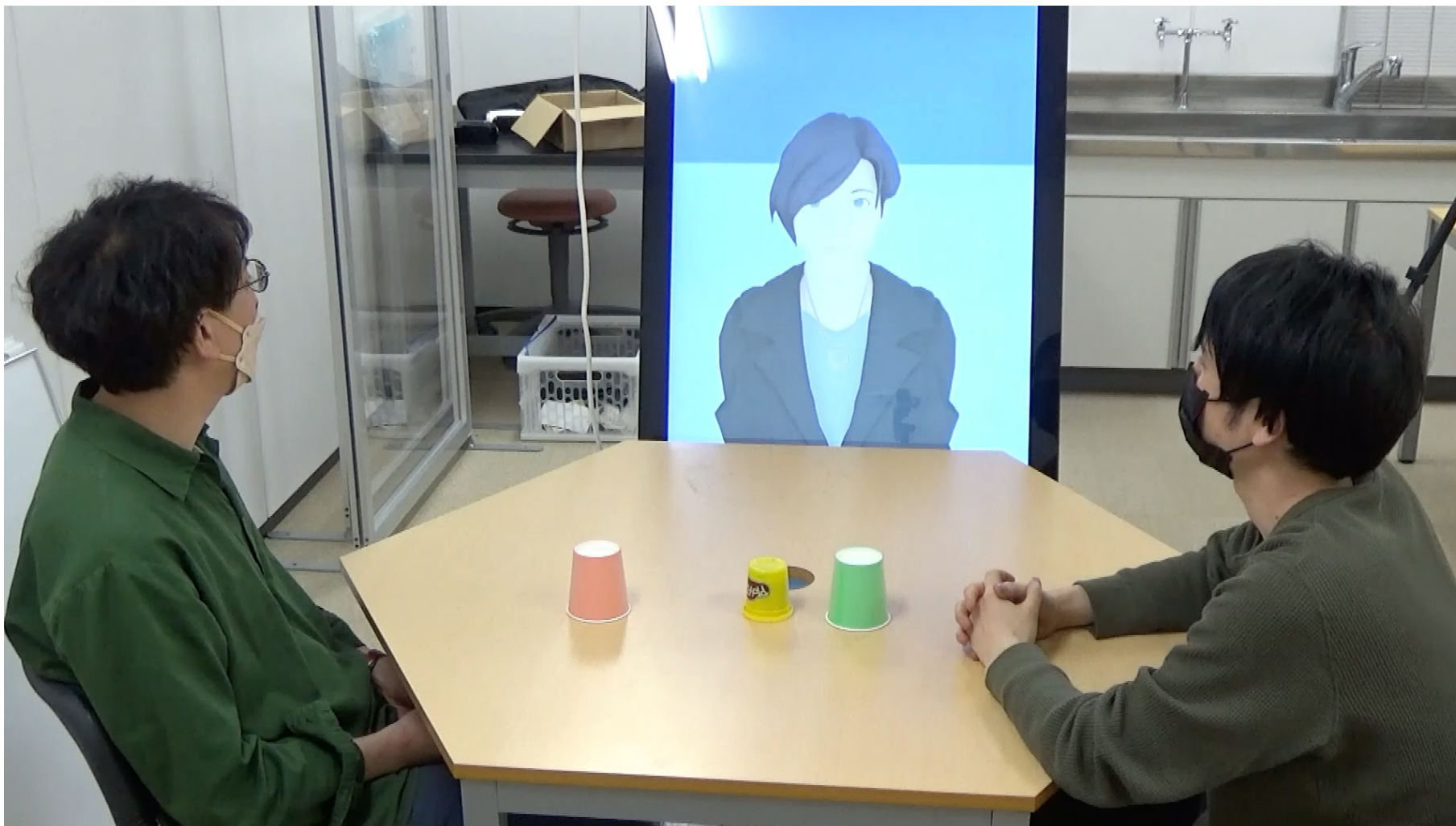
| 1. Appearance | 2. ToM video | 3. SST |

Greeting

A. High-level ToM
B. Low-level ToM

Statistical analysis

Statistical comparisons were made between groups in the three stages

Calculated the differences between the appearance and ToM video stages and the ToM video and SST stages according to the ToM levels

Comparisons between groups in the ToM video stage regarding the ToM levels were made by a Wilcoxon ranks sum test (H1)

Tested whether the distrust caused by the low level ToM video was weakened by the Wilcoxon rank-sum test (H2)
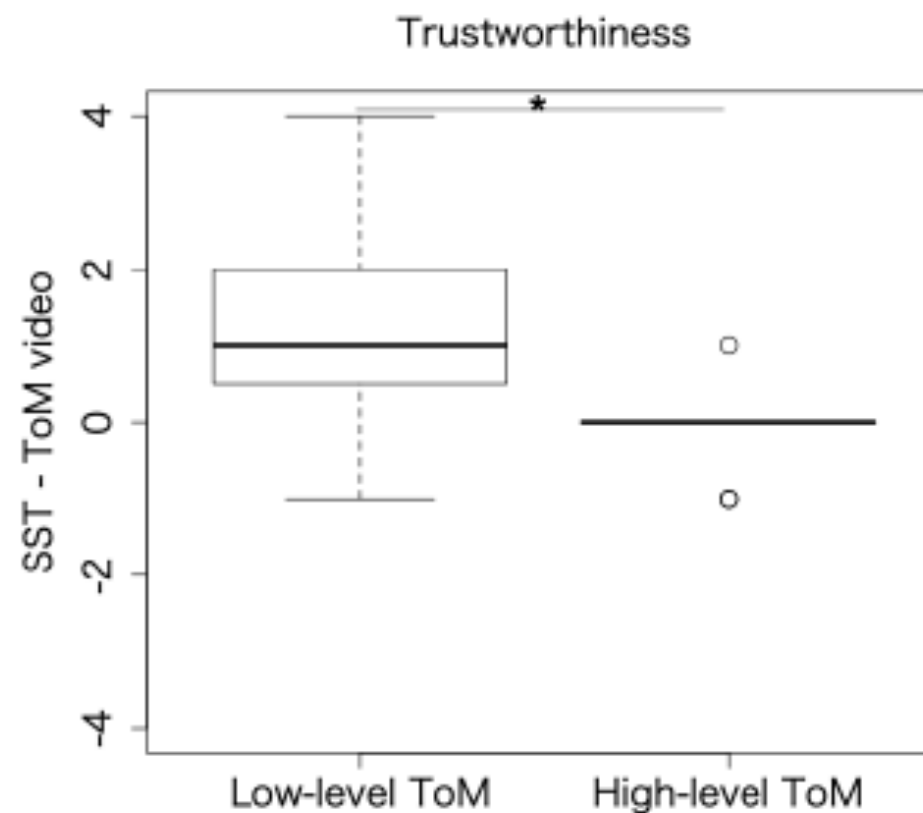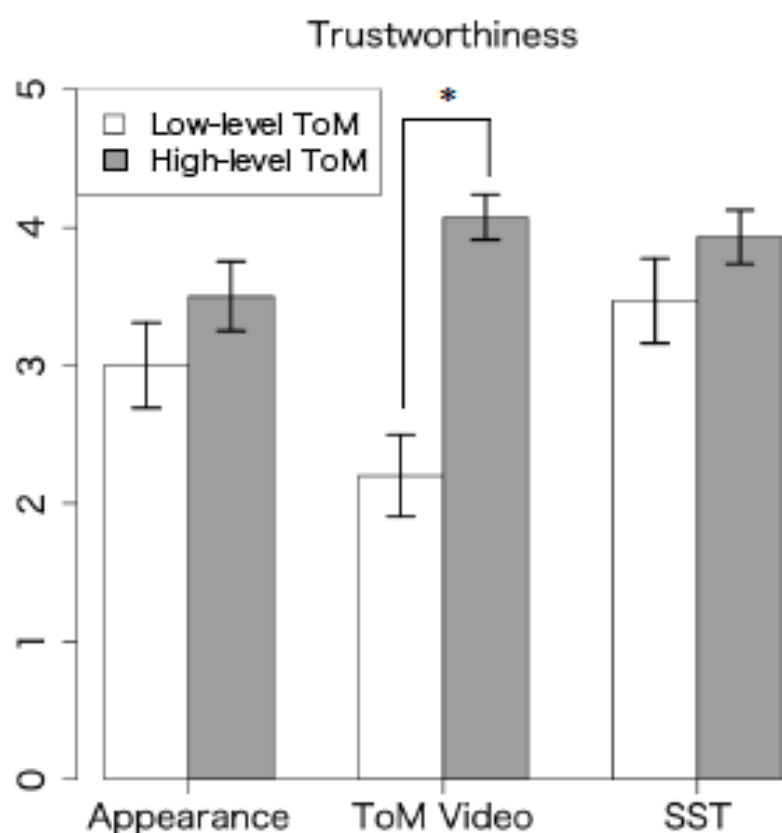
Results: acceptability

ToM levels affect acceptability as a trainer ($p<0.05$) (H1), and the SST enhances acceptability (H2)

TAPAS SUPPORTED BY ANR-CREST   2022/12/3
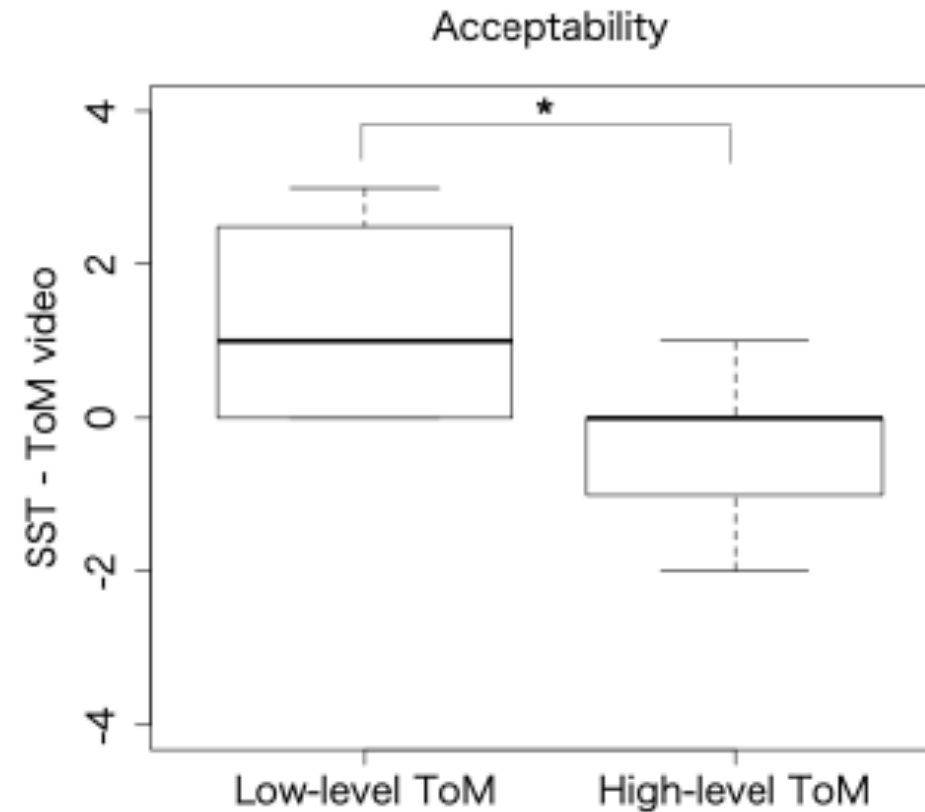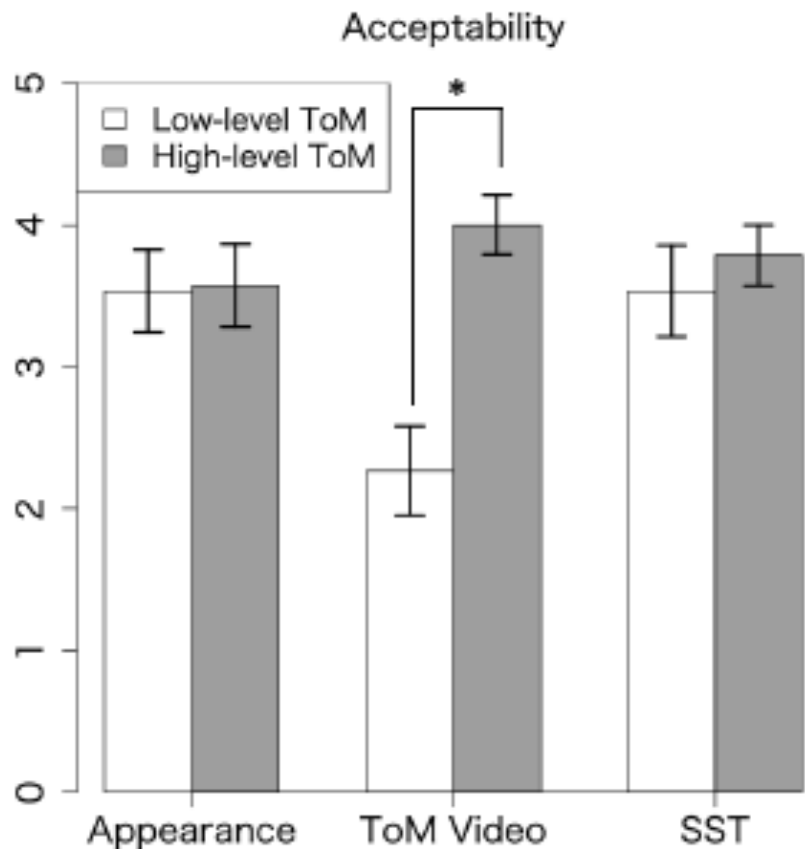
Results: trustworthiness

ToM levels affect trustworthiness as a trainer ($p<0.05$) (H1), and distrust was weakened by the SST process (H2)

Conclusions

Used the SST system and clarified the following two points of investigation

(H1) the virtual agents' ToM levels have effects on their acceptability and trust as trainers

(H2) distrust was weakened by the SST process

Future work

Integrating automatic ToM functions into the system

Since people with social difficulties struggle to understand ToM, incorporating such social cognition training into SST as role-playing is crucial

References

W. Mou, M. Ruocco, D. Zanatto and A. Cangelosi, "When Would You Trust a Robot? A Study on Trust and Theory of Mind in Human-Robot Interactions," 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 2020, pp. 956-962, doi: 10.1109/RO-MAN47096.2020.9223551.

Hiroki Tanaka, Satoshi Nakamura, The Acceptability of Virtual Characters as Social Skills Trainers: Usability Study, JMIR Human Factors, 9(1), 2022.

Hiroki Tanaka, Hidemi Iwasaka, Yasuhiro Matsuda, Kosuke Okazaki, Satoshi Nakamura, Analyzing Self-Efficacy and Summary Feedback in Automated Social Skills Training, IEEE Open Journal of Engineering in Medicine and Biology, vol.2, pp.65-70, 2021.

Takeshi Saga, Hiroki Tanaka, Yasuhiro Matsuda, Tsubasa Morimoto, Mitsuhiro Uratani, Kosuke Okazaki, Yuichiro Fujimoto, Satoshi Nakamura, Analysis of Feedback Contents and Estimation of Subjective Scores in Social Skills Training, 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), July 2022.

Thank you for your attention