Detecting Dementia and Enhancing Social Skills through Human and Embodied Agent Interaction

Satoshi Nakamura (NAIST)

with

Hiroki Tanaka (NAIST), Takashi Kudo (Osaka University),
and Kosuke Okazaki (Nara Medical University)

Two Topics

1. ## Automatic detection of early neurodegenerative disorder through interaction with the embodied agent. (-2018)

   - Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroaki Kazui, Takashi Kudo, Satoshi Nakamura, "Detecting Dementia through Interactive Computer Avatars, "IEEE Journal of Translational Engineering in Health and Medicine, 5(1), 1-11, Dec. 2017
   - Tsuyoki Ujiro, Hiroki Tanaka, Hiroyoshi Adachi, Hiroaki Kazui, Manabu Ikeda, Takashi Kudo, Satoshi Nakamura, "Detection of dementia from responses to atypical questions asked by embodied conversational agents," Proc. INTERSPEECH2018

2. ## Social skills training for the neurodevelopmental disorder through interaction with the embodied agent.

   - 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.

Two Topics

1. **Automatic detection of early neurodegenerative disorder through interaction with the embodied agent. (-2018)**

   - Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroaki Kazui, Takashi Kudo, Satoshi Nakamura, "Detecting Dementia through Interactive Computer Avatars, "IEEE Journal of Translational Engineering in Health and Medicine, 5(1), 1-11, Dec. 2017

   - Tsuyoki Ujiro, Hiroki Tanaka, Hiroyoshi Adachi, Hiroaki Kazui, Manabu Ikeda, Takashi Kudo, Satoshi Nakamura, "Detection of dementia from responses to atypical questions asked by embodied conversational agents," Proc. INTERSPEECH2018

2. **Social skill training for the neurodevelopmental disorder through interaction with the embodied agent.**

   - 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.

Aim

Aim

Early detection of dementia (Don't need to go to hospital!)

Proposal

- During interaction with agents
- Tools that can be used repeatedly for daily use

Detection of dementia from responses to *atypical questions*

Suspected of dementia

No suspected of dementia

Neuropsychological Tests

Typical question: Based on neuropsychological tests

○ Frequently used screening for dementia

  ◦ MMSE [Folstein et al. 1975], "What's the date today?"

  ◦ WMS-R [Wechsler 1997]

  → **Experts are necessary for these tests**

Related works

  ◦ Roark et al. 2011:   Analyze response during Wechsler memory recall task.
  ◦ Aramaki et al. 2016:   Find a relationship between cognitive and language abilities during HDS-R.
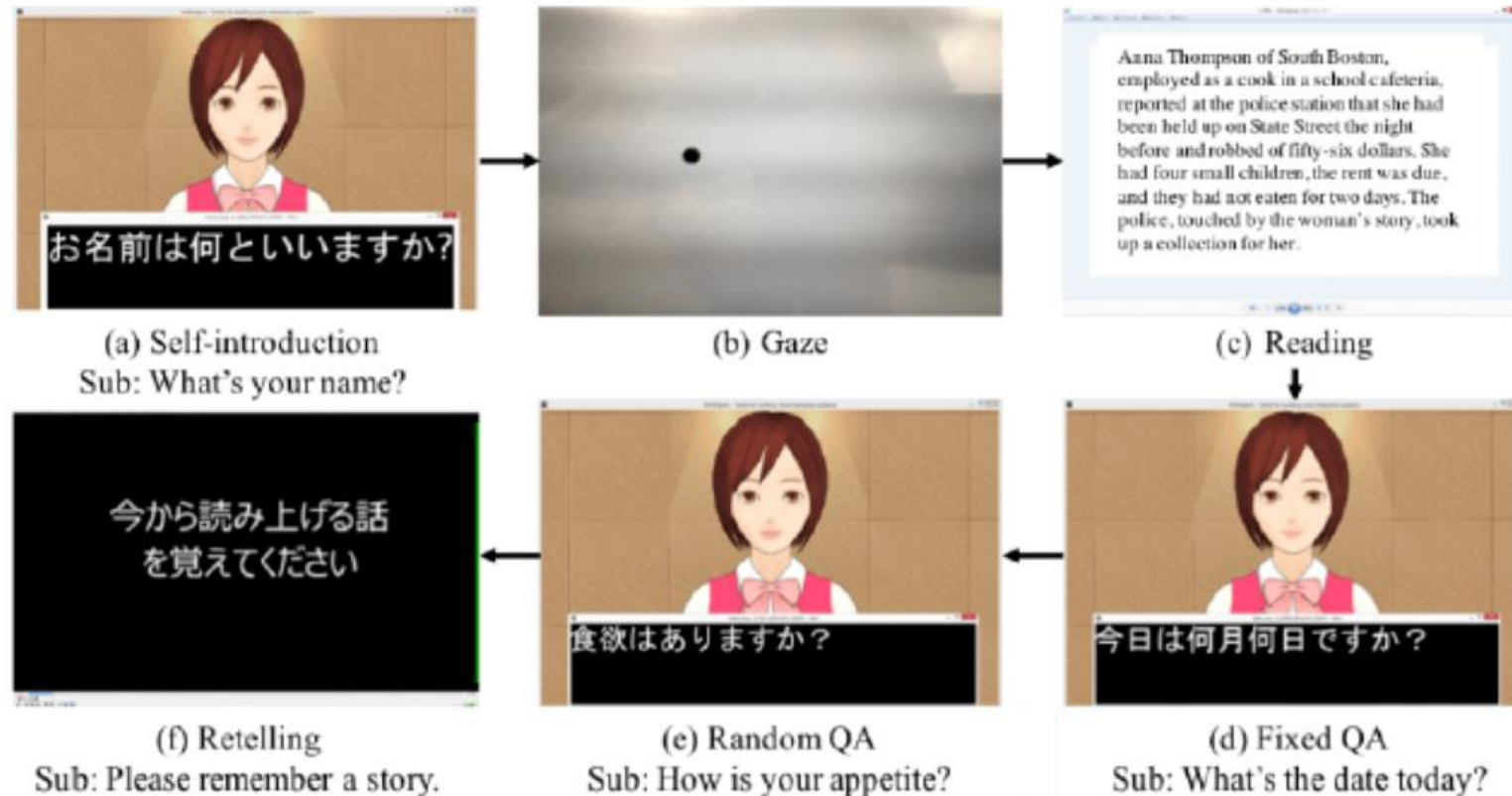 ◦ Tanaka et al. 2017:   Analyze speech and language features during MMSE test.

■ Possibilities of detecting dementia from speech and language features
■ These studies are not suitable for daily use

Atypical question: Not based on neuropsychological test

Methods

10 minuets to complete all tasks

(a) Self-introduction
Sub: What's your name?

(b) Gaze

(c) Reading

(f) Retelling
Sub: Please remember a story.

(e) Random QA
Sub: How is your appetite?

(d) Fixed QA
Sub: What's the date today?

[Ujiro et al., 2018]   [Tanaka et al., 2017]

Questions sets

Fixed QA: The system asks three fixed queries based on MMSE:

1) What's the date today? (Q1)
2) Tell me something interesting about yourself.(Q2)
3) How did you come here today? (Q3)

Randomly select 5 from 13 questions

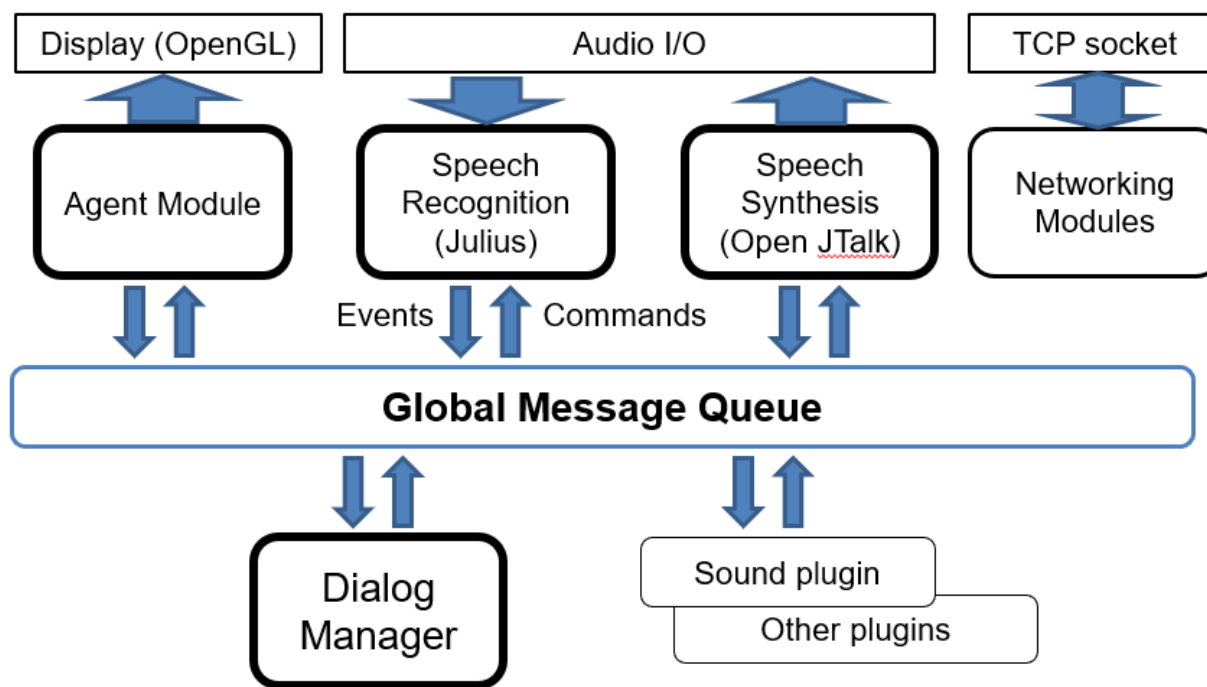|	Content
Q1	Please tell me about your family.
Q2	Please tell me something that you feel is stressful in your life.
Q3	What is your hobby?
Q4	What is your favorite song?
Q5	Please tell me about Yujiro Ishihara.
Q6	Please tell me about Shigeo Nagashima.
Q7	Please tell me about Hibari Misora.
Q8	Who is Japan's Prime Minister?
Q9	What season is it now?
Q10	What year is it?
Q11	Are you left-handed or right-handed?
Q12	Do you sleep well?
Q13	How is your appetite?

Retelling

■ The system reads aloud from a different part of the Wechsler logical memory in the WMSR [14], and users paraphrase the passage.

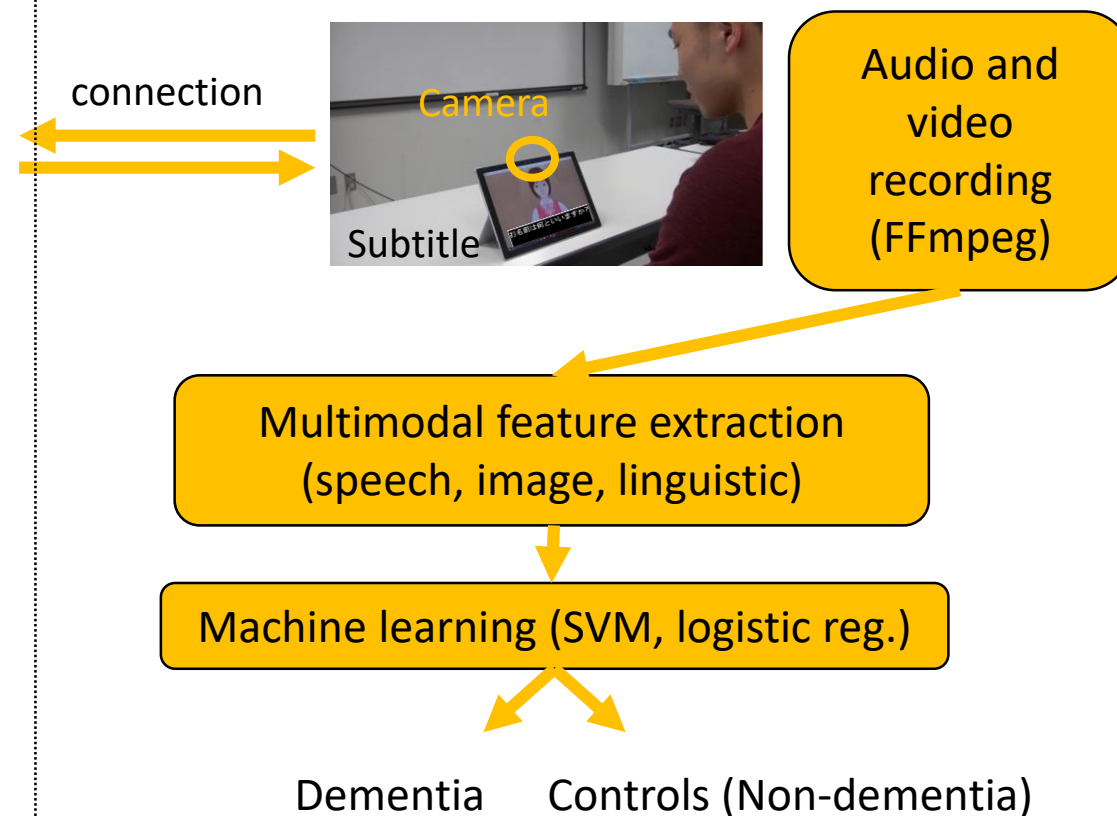■ Record the audio sound spoken by a clinical psychologist in advance. The system plays the sound.

2023/6/23

Embodied Agent System

MMD system architecture

https://mmdagent-ex.dev/docs/system/

Dementia detection module

connection

Audio and video recording (FFmpeg)

Multimodal feature extraction (speech, image, linguistic)

Machine learning (SVM, logistic reg.)

Dementia   Controls (Non-dementia)

Experiments: Participants

Group	N	Age	MMSE	Education
		M (SD)	M (SD)	M (SD)
Non-dementia	12	74.5 (4.3)	27.5 (1.8)	8.8 (2.6)
Dementia	12	75.9 (7.3)	21.2 (5.1)	13.9 (3.8)

Nine Alzheimer's disease (AD), one normal pressure hydrocephalus (NPH), one MCI, and one AD+NPH. Per the existing criteria, the dementia group was diagnosed by the psychiatrists of the Osaka University Medical School-affiliated hospital based on DSM-IV-TR [4]. We obtained the age, MMSE score, and educational history of all participants.

Features
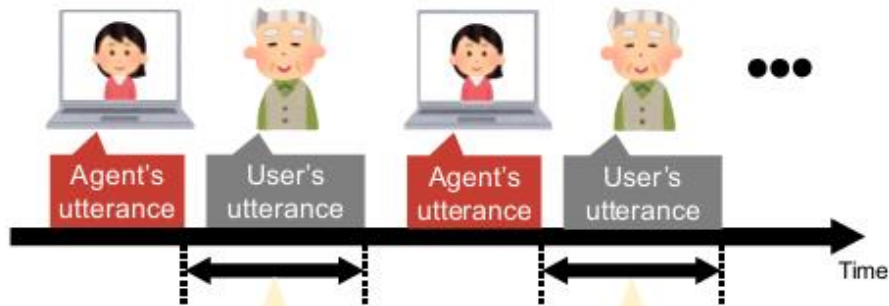
● Speech Features: Snack sound tool kit

  ●F0, Power, Voice quality

●Gap features

  ●Time delay from the end of the agent utterance to the beginning of the user's response.

●Language features

  ●Type Token Ratio (TTR)

  ●Simpson's D value

$$F0_{cov} = \frac{sd(f0)}{mean(f0)}.$$

Normalized F0 variation per utterance

$$h1a3 = h1 - a3.$$

Breathiness: 1-st harmonic amplitude -3rd formant amplitude

$$Gap = t_a - t_q.$$

Gap: the start time of the user response – the end time of the agent's question

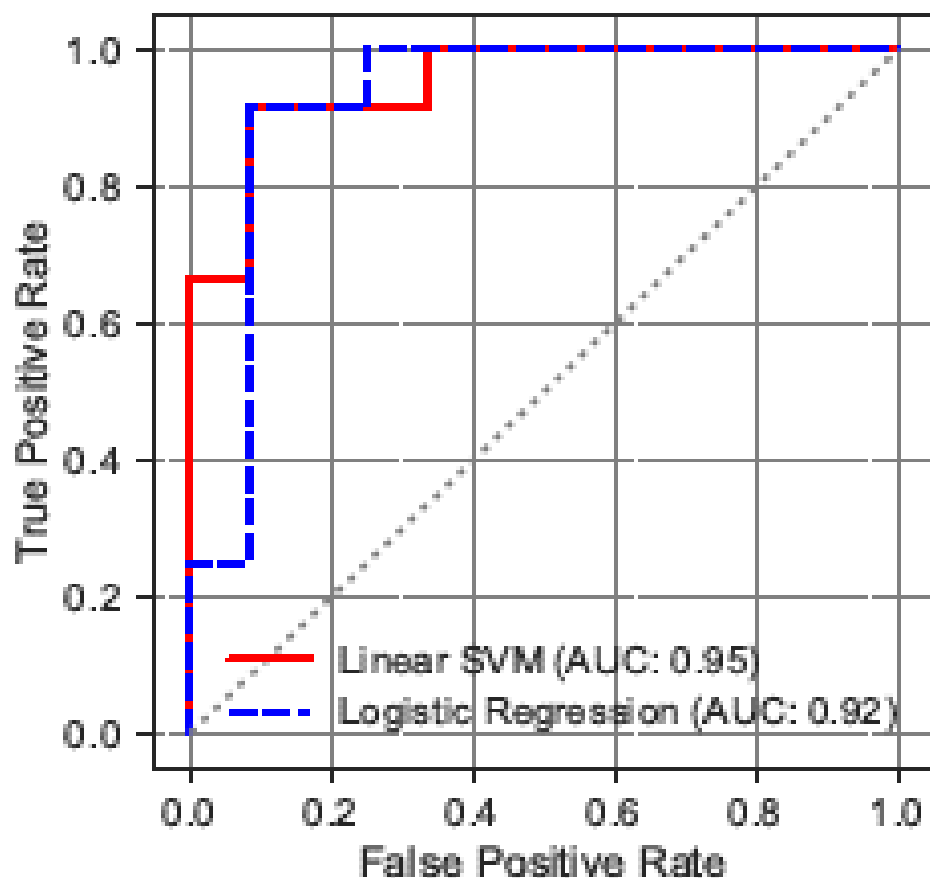$$TTR = \frac{\sum\limits_{all\ m} V(m, N)}{\sum\limits_{all\ m} mV(m, N)}.$$

Ratio of the total vocabulary to the overall words

$$D = \sum\limits_{all\ m} V(m, N)\frac{m}{N}\frac{m-1}{N-1}.$$

Normalized for the sample size

Results: AUC for Randomly selected questions

Input (classification model)	AUC	Accuracy	
Gap only (SVM)	0.69	0.63	
MMSE (SVM)	**0.85**	**0.83**	← Frequently used neuropsychological tests
Speech features (SVM)	0.85	0.83	
Speech features (LR)	0.90	0.83	
Speech and language features (LR)	0.92	0.92	
Speech, language and image features (SVM)	**0.93**	**0.83**	← Previous work [Tanaka et al. 2017]
Speech and language features (SVM)	**0.95**	**0.92**	← **This work!**

Normalized the features (Mean: 0, SD: 1)
Model evaluation: Leave-one-participant-out, ROC Curve

Demo

Two Topics

1. Automatic detection of early neurodegenerative disorder through interaction with the embodied agent.

   ◦ Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroaki Kazui, Takashi Kudo, Satoshi Nakamura, "Detecting Dementia through Interactive Computer Avatars, "IEEE Journal of Translational Engineering in Health and Medicine, 5(1), 1-11, Dec. 2017

   ◦ Tsuyoki Ujiro, Hiroki Tanaka, Hiroyoshi Adachi, Hiroaki Kazui, Manabu Ikeda, Takashi Kudo, Satoshi Nakamura, "Detection of dementia from responses to atypical questions asked by embodied conversational agents," Proc. INTERSPEECH2018

2. Social skills training for the neurodevelopmental disorder through interaction with the embodied agent.

   ◦ 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.
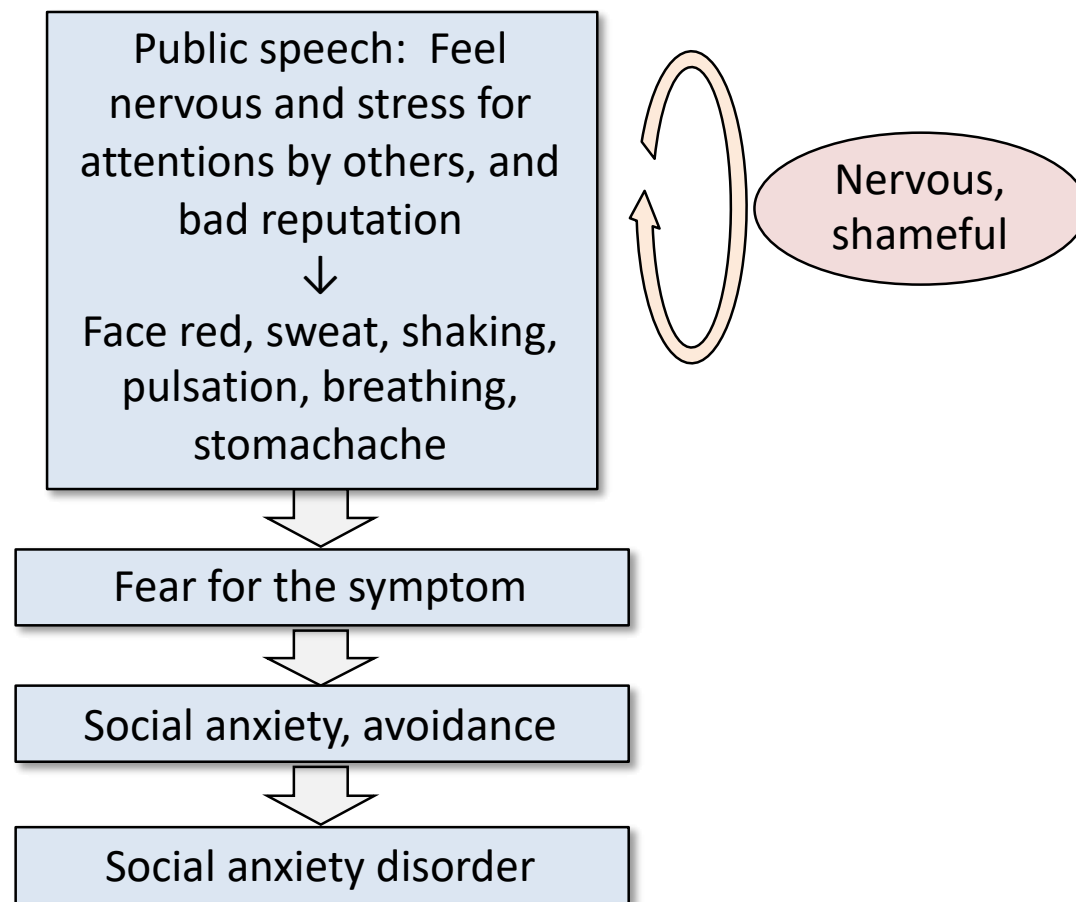
Training Adapted Personalized Affective Social Skills with Cultural Virtual Agents (TAPAS)

Problem: Social anxiety in schools and workplace

Approach: verbal/non-verbal interactive training system **by embodied conversational agent (ECA)**

- ◦ Target：
  General populations, Social Anxiety Disorder (SAD)
  Autism Spectrum Disorder (ASD), Schizophrenia (SZ)

- ◦ **Behavioral training**：
  Social skills training（SST）
- ◦ Cognitive training：
  Cognitive Behavioral Therapy（CBT）

Public speech: Feel nervous and stress for attentions by others, and bad reputation
↓
Face red, sweat, shaking, pulsation, breathing, stomachache

Nervous, shameful

Fear for the symptom

Social anxiety, avoidance

Social anxiety disorder

Training **A**dapted **P**ersonalized **A**ffective **S**ocial Skills with Cultural Virtual Agents: ANR-CREST Project

Social Skills Training

- Social Skills Training by Alan Bellack*

  - Social skills training (SST) is a type of <u>behavioral therapy</u> for people with mental disorders or developmental disabilities.
  - SST may be used by psychiatrists, teachers, therapists, or other professionals

*Bellack, A. S.: Social skills training for schizophrenia: A step-by-step guide, Guilford Press (2004)
https://www.verywellmind.com/social-skills-4157216

SATOSHI NAKAMURA@ NAIST, ANR-CREST

2023/6/23

Virtual agents and robots for social skills training (SST)

Robins+1998

Kumazaki+2018

Better communication
with Virtual agents!
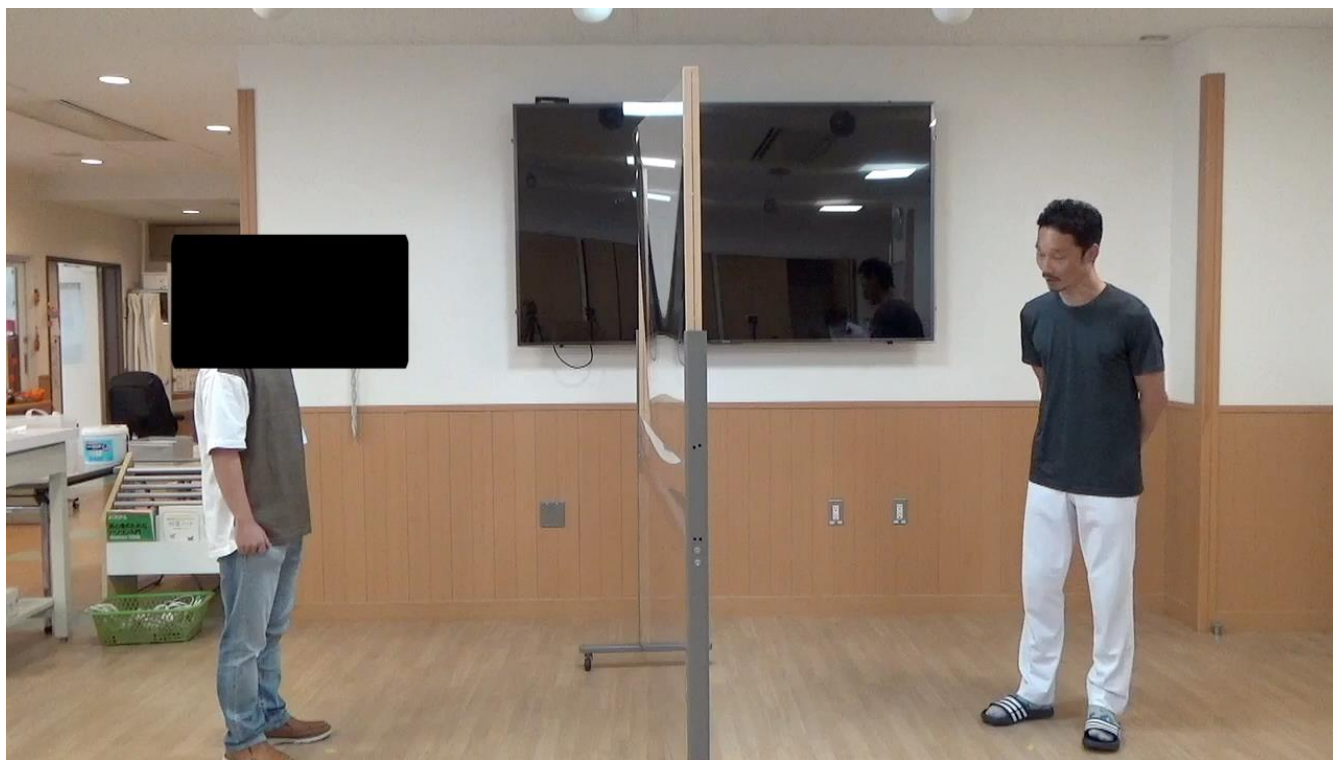

Hoque+2015

Nijman+2020

SST data collection to generate human-like feedback

Adults: Schizophrenia (N=16), Autism spectrum disorder (N=15), Healthy controls (N=21)
Children: Autism spectrum disorder (N=16), Healthy controls (N=17)

Assessments (ADOS, SRS2, trainers' ratings etc), eye gaze
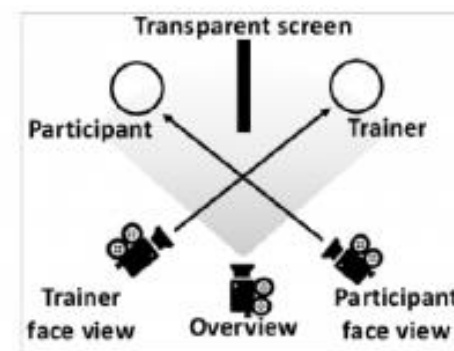

2*Kinect
2*Face
1*whole

Behavioral measurements (basic module for SST)

Multimodal prediction model for Social Responsiveness Scale-2 (SRS2), N=27

Audio speech, Facial action units, and BERT similarity in speaking [Saga+2022]

Correlation coefficient for SRS2 = 0.76

1	Feel comfortable to stay with others compared to stay alone	1,2,3,4
2	My facial expression makes misunderstanding to my real feeling.	1,2,3,4
3	I am confident to communicate to others	1,2,3,4

1) Calculate embeddings

2) Calculate cosine similarity
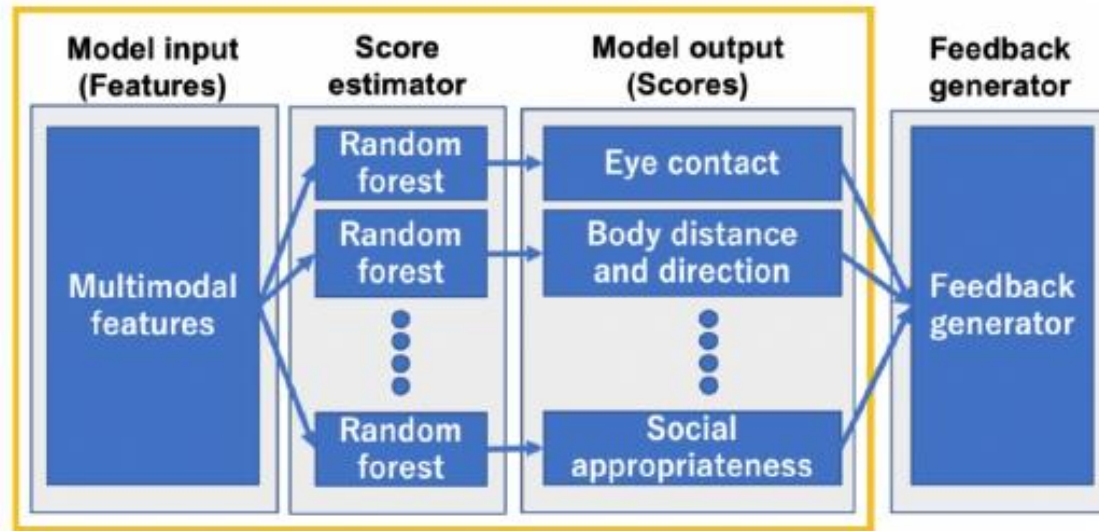
3) Take average of 2)
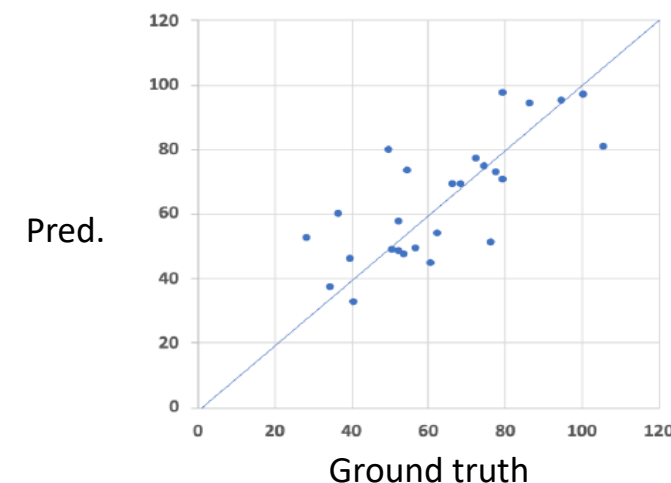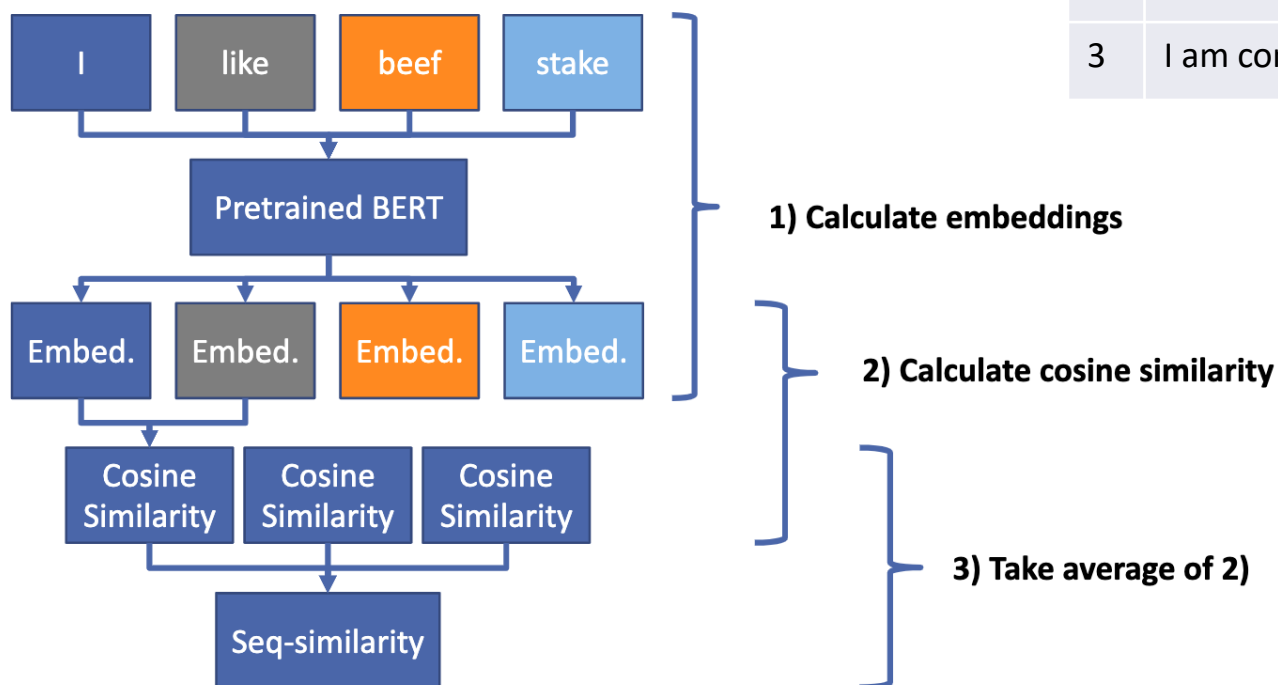
Pred.

Ground truth

Feedback generation using prediction model

RQ: Can multimodal features predict trainer's score? [Saga+, 2022]

Applied the behavioral measurements model to the collected data

Four SST tasks: ASK, DECLINE, TELL, LISTEN

Ground truth evaluation: 1 to 5 for 7 skills by two third-party trainers

Prediction results indicating correlation coefficient for third-party evaluation scores

Skill Label	ASK	DECLINE	TELL	LISTEN
Eye contact	0.04	**0.53**	0.16	**0.48**
Body direct. and dist.	0.16	0.14	0.33	**0.30**
Facial expression	**0.29**	**0.44**	**0.63**	**0.33**
Voice variation	**0.36**	**0.42**	**0.44**	**0.31**
Clarity	**0.42**	**0.44**	0.16	0.07
Fluency	0.17	**0.36**	0.19	**0.36**
Social Appropriateness	**0.48**	**0.50**	0.27	0.18

Feedback generation

The feedback is generated by template sentences according to the highest and lowest scores

Reason and advice were made as referring to human SST data and finalized by consultation with NMU
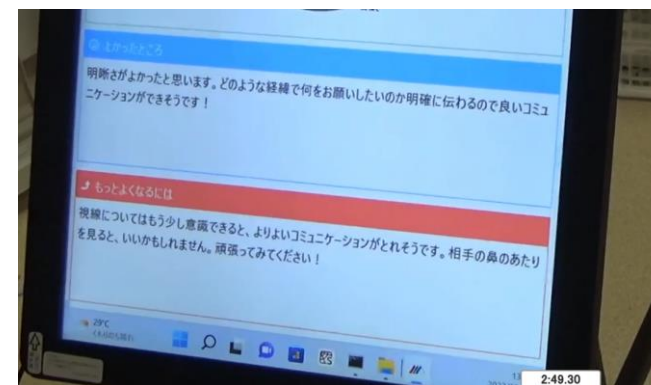
Template example

Highest score: Clarity

Your [Skill name] was good. Since [Reason], I think you can do great communication!

clarity

they can understand what you want clearly
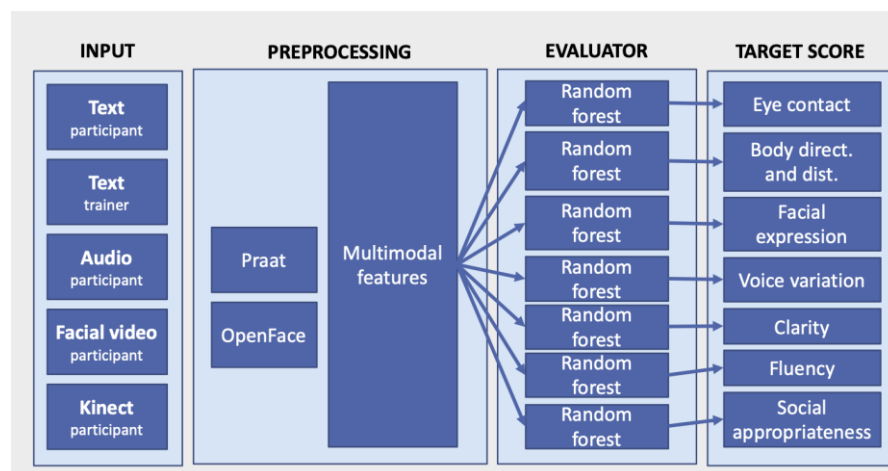
Lowest score: Eye contact

It seems if you can be more aware of [Skill name], I think you can take great communication. I think you better to [Advice] for improvement. Hang in there!
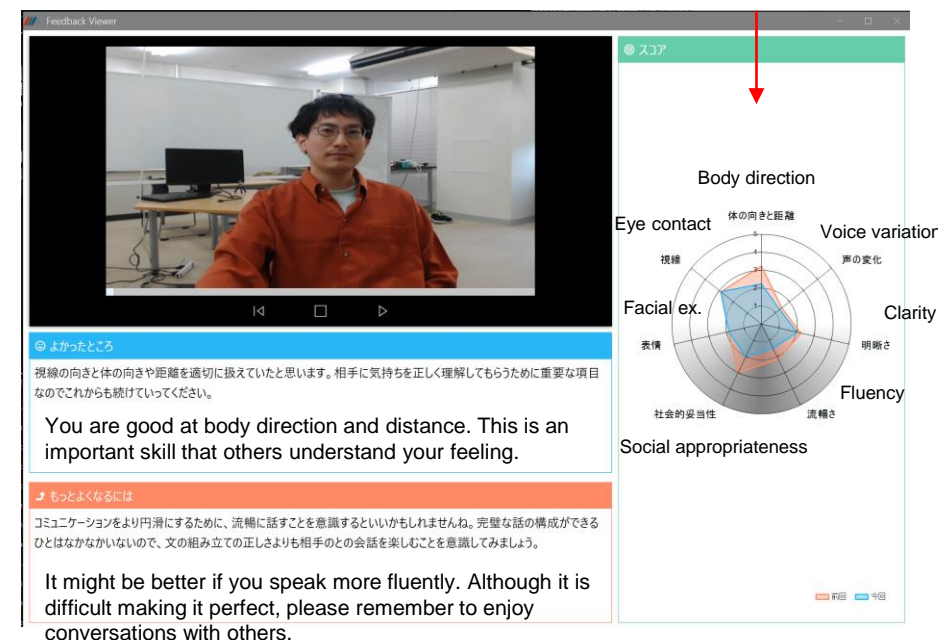
watch around their nose

eye contact

SST system

Roleplay w/ agents ⟶ Prediction ⟶ Feedback

Predicted scores

Demo video

NAIST, ANR-CREST

Summary

1. Automatic detection of early dementia through interaction with the embodied agent.
   - Audio, language, image features + Gap features (AUC: 0.95 Acc: 0.93)
   - Atypical questions can be used for daily diagnoses of early dementia

2. Social skills training for the neurodevelopmental disorder through interaction with the embodied agent.
   - Social skills training as a behavioral training for communication disabilities: Anxiety disorder, Autism spectrum disorder, Schizophrenia
   - Social skills training by an embodied computational agent
   - Sensing multimodal features and automatic summary feedback

Members

Early dementia detection:
- Hiroki Tanaka, Tsuyoki Ujiro (NAIST)
- Takashi Kudo, Hiroyoshi Adachi, Hiroaki Kazui, Manabu Ikeda (Osaka University)

Social Skills Training: (TAPAS project)
- Hiroki Tanaka, Takeshi Saga, Yuichiro Fujimoto (NAIST)
- Kosuke Okazaki, Yasuhiro Matsuda, Tsubasa Morimoto, Mitsuhiro Uratani (Nara Medical University)
- Takashi Kudo (Osaka University)
- Catherine Pelachaud（France CNRS-ISIR)
- Jean Claude, Martin (France CNRS-LISN）

Reference

Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Takashi Kudo, Satoshi Nakamura, Automatic Detection of Very Early Stage of Dementia through Multimodal Interaction with Computer Avatars, ACM International Conference on Multimodal Interaction (ICMI), pp.261-265, Nov. 2016.

Hiroki Tanaka, Hiroyoshi Adachi, Norimichi Ukita, Manabu Ikeda, Hiroaki Kazui, Takashi Kudo, Satoshi Nakamura, Detecting Dementia through Interactive Computer Avatars, IEEE Journal of Translational Engineering in Health and Medicine, vol.5(1), pp.1-11, Dec. 2017.

Tsuyoki Ujiro, Hiroki Tanaka, Hiroyoshi Adachi, Hiroaki Kazui, Manabu Ikeda, Takashi Kudo, Satoshi Nakamura, Detection of dementia from responses to atypical questions asked by embodied conversational agents, Interspeech, pp.1691-1695, Sep. 2018.

[Saga+2022] Takeshi Saga, Hiroki Tanaka, Hidemi Iwasaka, Satoshi Nakamura, Multimodal Prediction of Social Responsiveness Score with BERT-based Text Features, IEICE Trans. Information and Systems, 2022.

[Tanaka+2022] Hiroki Tanaka, Satoshi Nakamura, Acceptability of Virtual Characters as a Social Skills Trainer: Usability Study, JMIR Human Factors, 2022.

[Saga+2021] Takeshi Saga, Hiroki Tanaka, Hidemi Iwasaka, Yasuhiro Matsuda, Tsubasa Morimoto, Mitsuhiro Uratani, Kosuke Okazaki, Yuichiro Fujimoto and Satoshi Nakamura, Multimodal Dataset of Social Skills Training in Natural Conversational Setting, ICMI 2021 Workshop: 2nd Workshop on Social Affective Multimodal Interaction for Health, pp.395-399, Oct. 2021.

[Saga+2022] 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, EMBC2022.

[Tanaka+2021] 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

[Tanaka+2020] Hiroki Tanaka, Hidemi Iwasaka, Hideki Negoro, Satoshi Nakamura, Analysis of Conversational Listening Skills toward Agent-based Social Skills Training, Journal on Multimodal User Interfaces, volume 14, issue 1, pp.73-82, 2020.

[Tanaka+2017] Hiroki Tanaka, Hideki Negoro, Hidemi Iwasaka, Satoshi Nakamura, Embodied Conversational Agents for Multimodal Automated Social Skills Training in People with Autism Spectrum Disorders, PloS One, vol.12(8), pp.1-15, Aug. 2017.

[Tanaka+2016] Hiroki Tanaka, Sakriani Sakti, Graham Neubig, Tomoki Toda, Hideki Negoro, Hidemi Iwasaka, Satoshi Nakamura, Teaching Social Communication Skills through Human-Agent Interaction, ACM Transactions on Interactive Intelligent Systems, vol.6(2), pp.1-26, 2016.

Thank you for listening!