ANR AGENCE NATIONALE DE LA RECHERCHE

JST Japan Science and Technology Agency CREST

TAPAS: Training Adapted Personalised Affective Social Skills with Cultural Virtual Agents

Satoshi Nakamura, Nara Institute of Science and Technology

Hiroki Tanaka, Nara Institute of Science and Technology

Hidemi Iwasaka, Nara Medical University

Takashi Kudo, Osaka University

ANR AGENCE NATIONALE DE LA RECHERCHE

Limsi

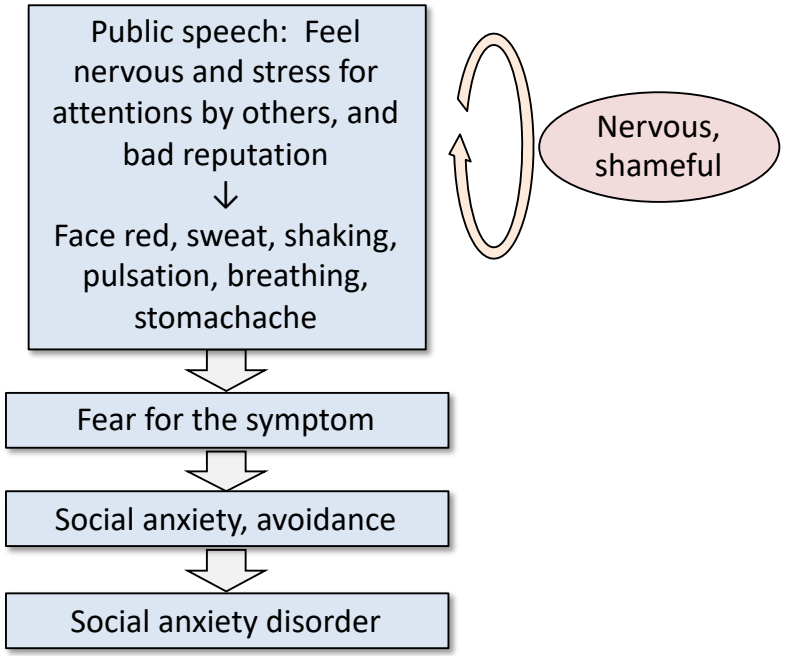
ISIR INSTITUT DES SYSTÈMES INTELLIGENTS ET DE ROBOTIQUE

TAPAS: Background

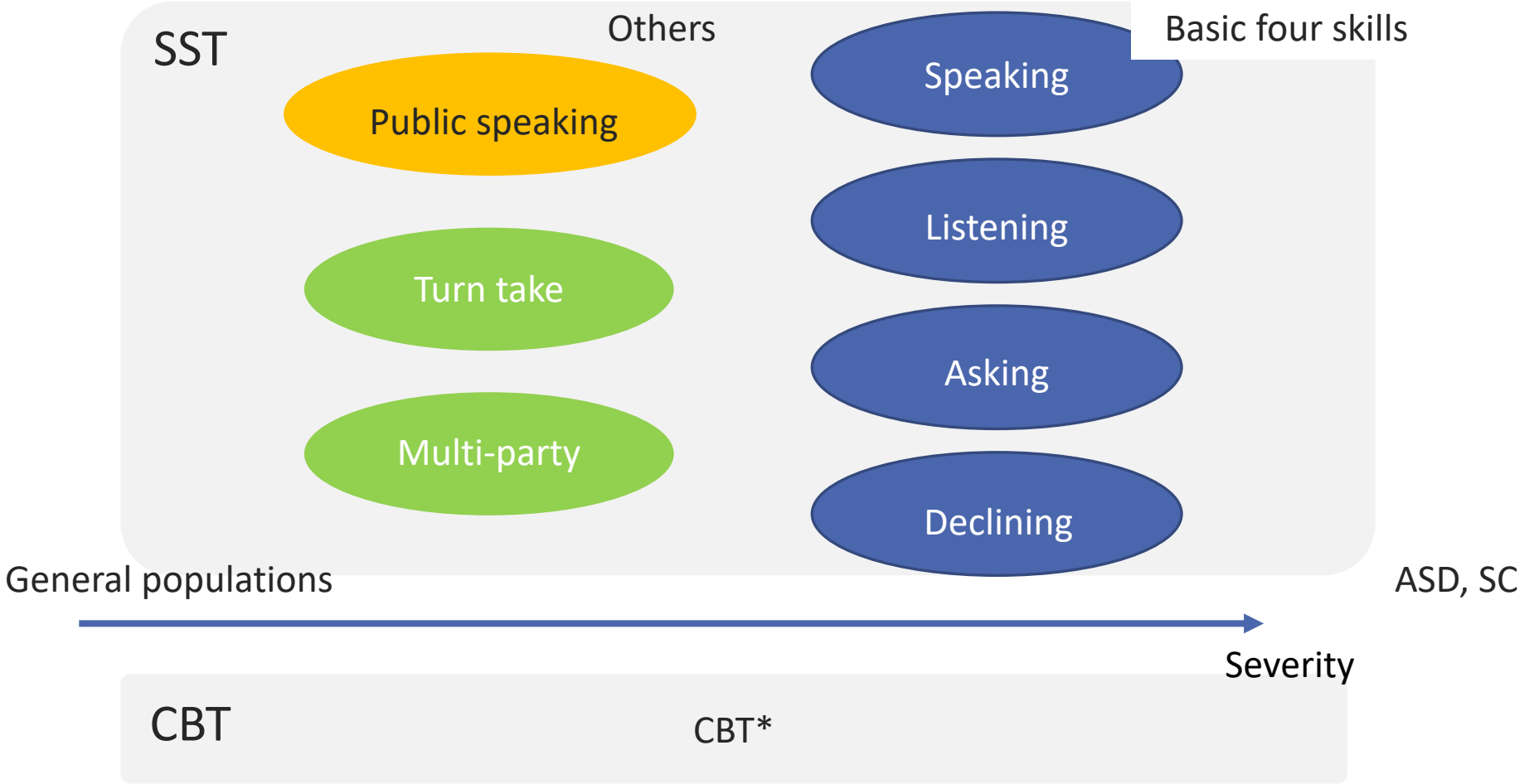
Increase of people who are not good at social communication and social anxiety disorder in schools and workplace (3 to 13%)

Try to solve by verbal / non-verbal interactive training system by Embodied Conversational Agent (ECA)

- Target：general populations、Autism（ASD: Autism Spectrum Disorder）、Schizophrenia, SAD: Social Anxiety Disorder
- Behavioral training：Apply Social skills training（SST）
- Cognitive training：Apply Cognitive Behavioral Training（CBT）



SST skills and Target populations



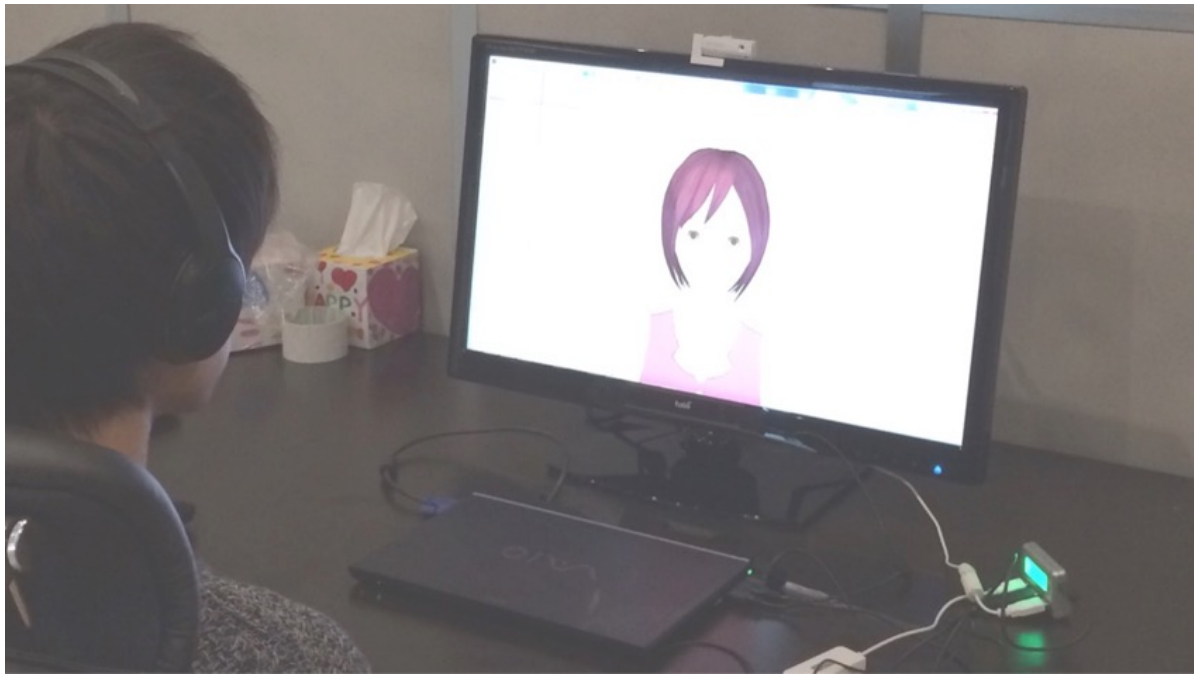
*CBT is used independent to severity

Automated social skills trainer

Training tools for social communication difficulties

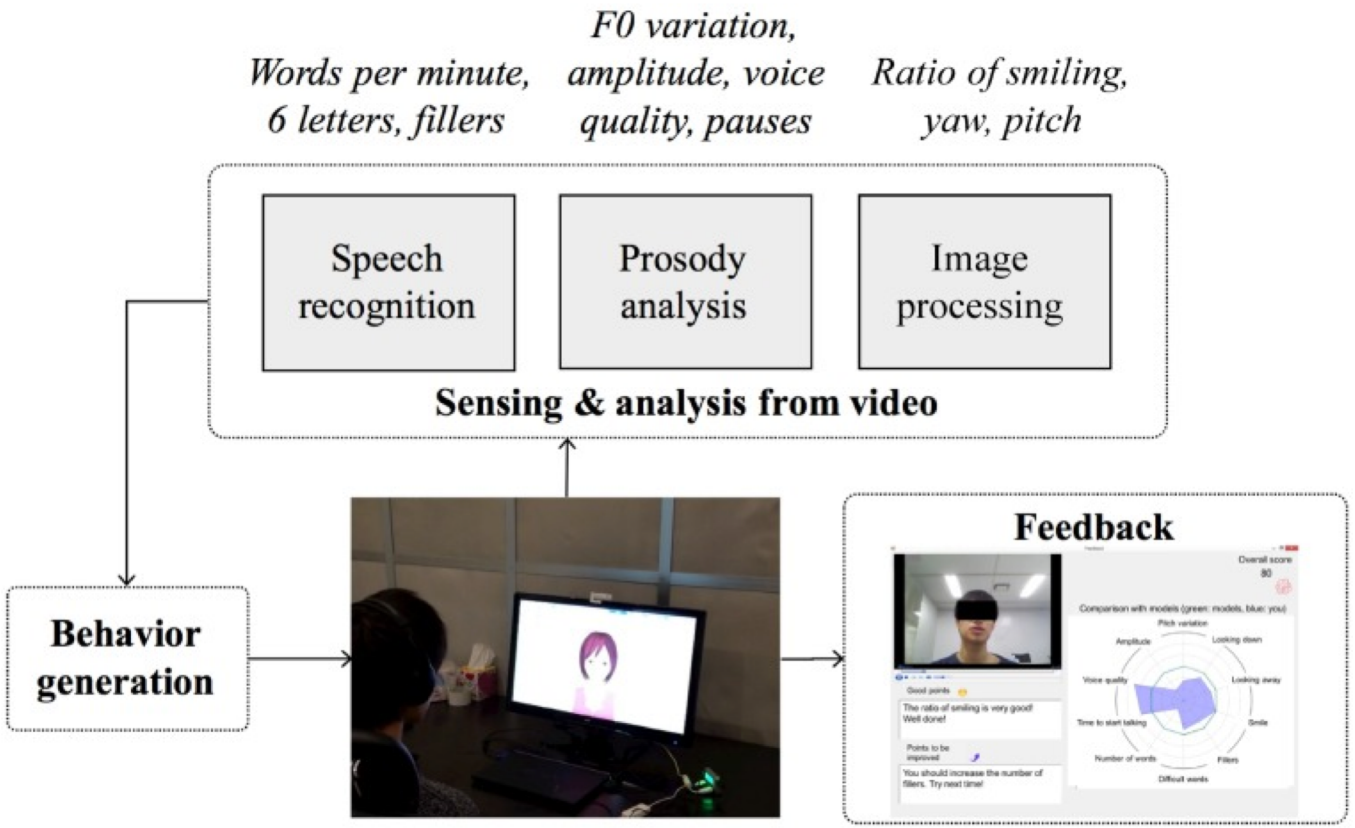
People with autism have good systemizing skills [Baron-Cohen et a., 2009]

Target skill: speaking skills [Tanaka, et al., 2015] and listening skills [Tanaka, et al., 2020]



Role-play

Analyze behaviors of users and generate agent's actions

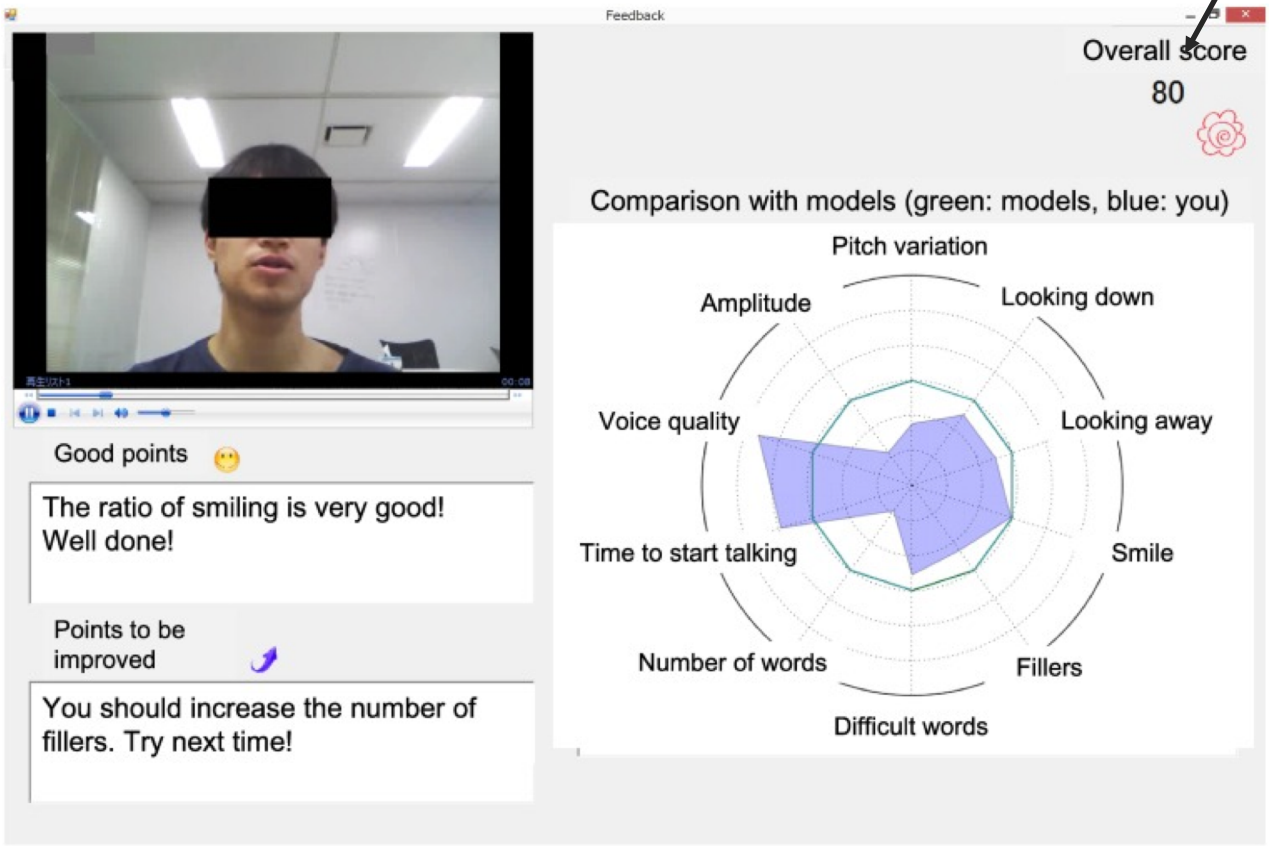


Features based on [Tanaka et al., 2014]

Summary feedback

After role-play, the system displays feedback

Repetitive training until mastery

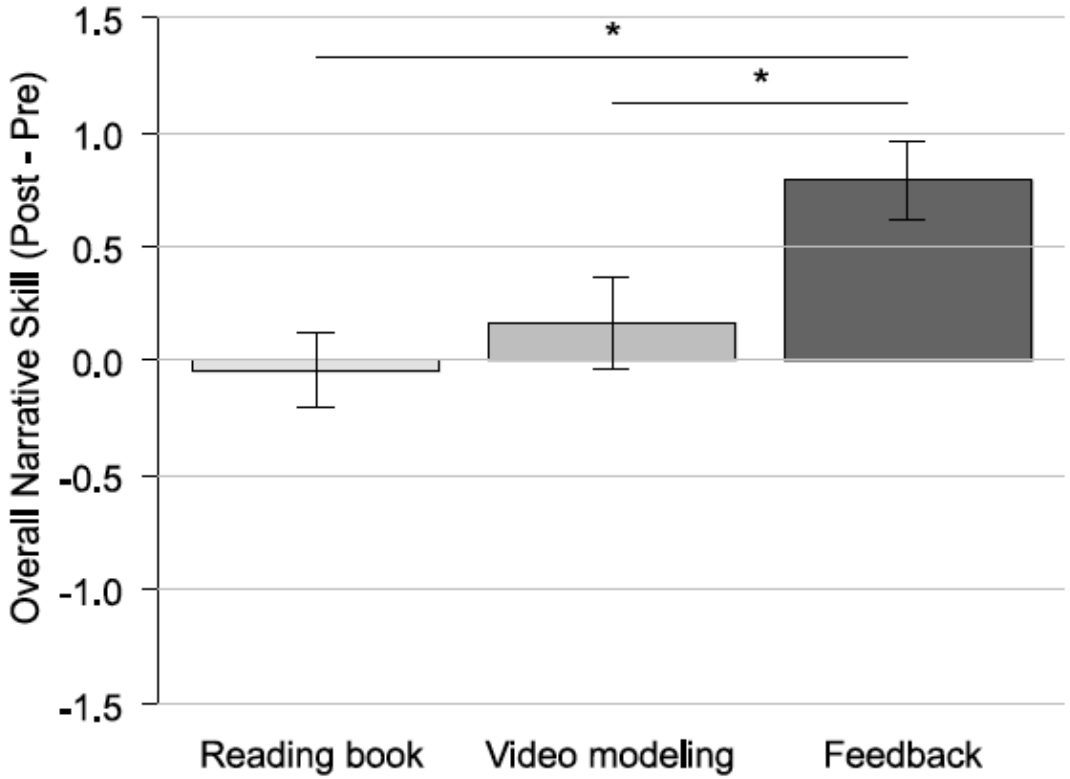


Multiple Linear regression on subjective ratings
(input: words per minute, amplitude, words over six letters, and smiling ratio)

Training effect

Overall narrative skill: 17 subjects, average of two students' ratings (1 to 7)

One-way ANOVA: ($F[2,24]=4.70$, $p<.05$) , Post hoc comparisons with Bonferroni's correction ($p<0.05$)

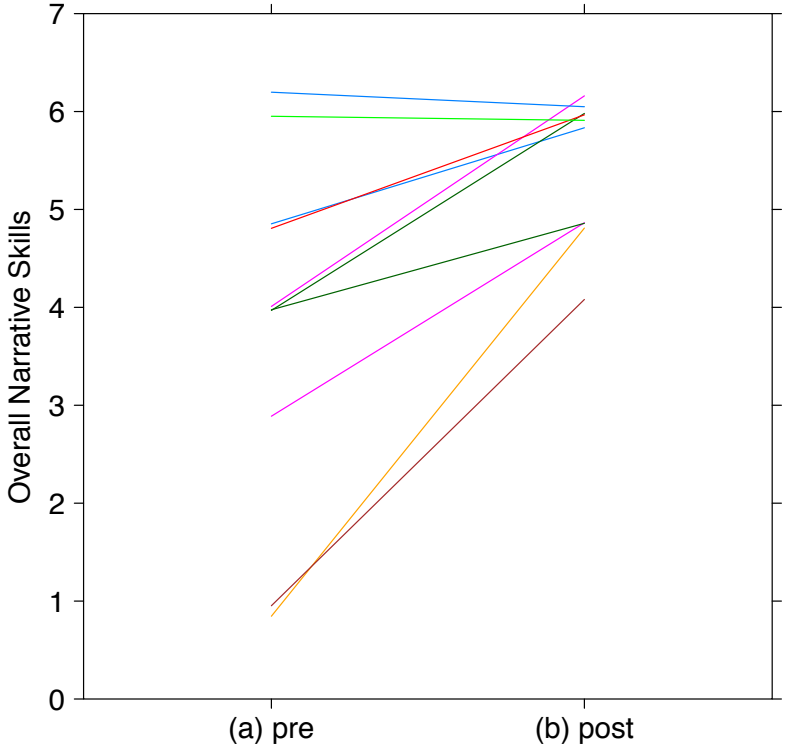


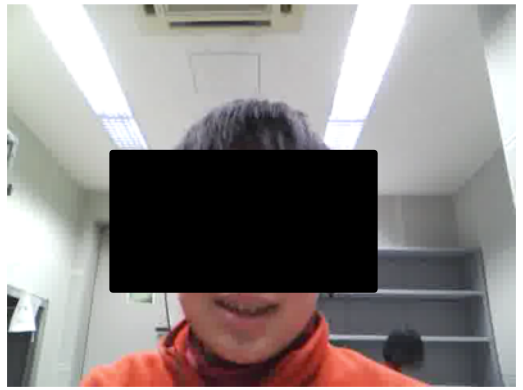
Children with autism spectrum disorders

Communicate to unfamiliar person (pre and post)

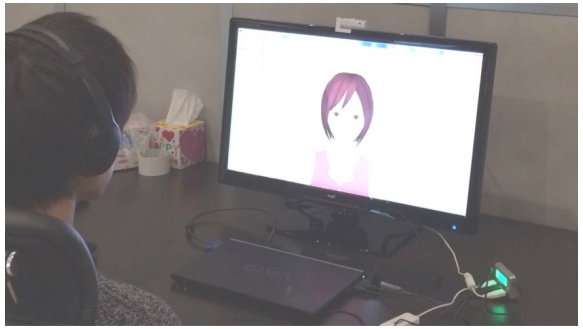
Overall narrative skill: clinical psychologist's rating (1 to 7)

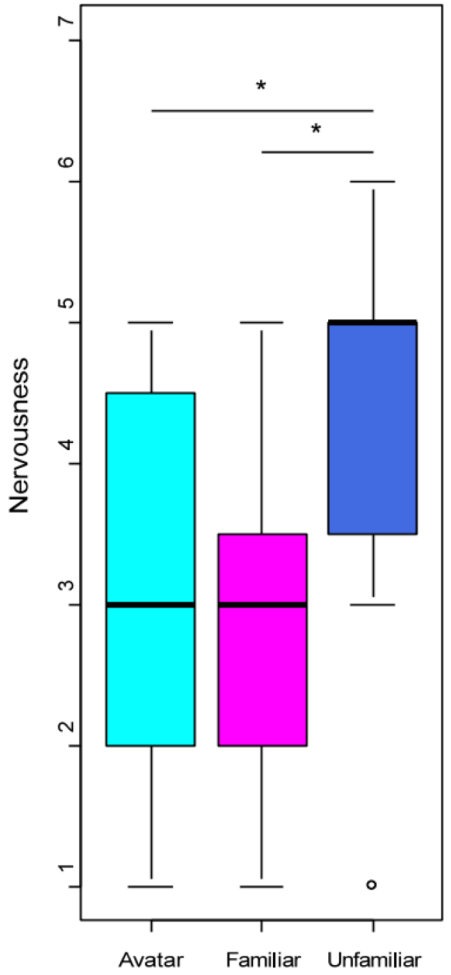
Pre-post comparison ($p = 0.002$, Cohen's $d = 1.17$)

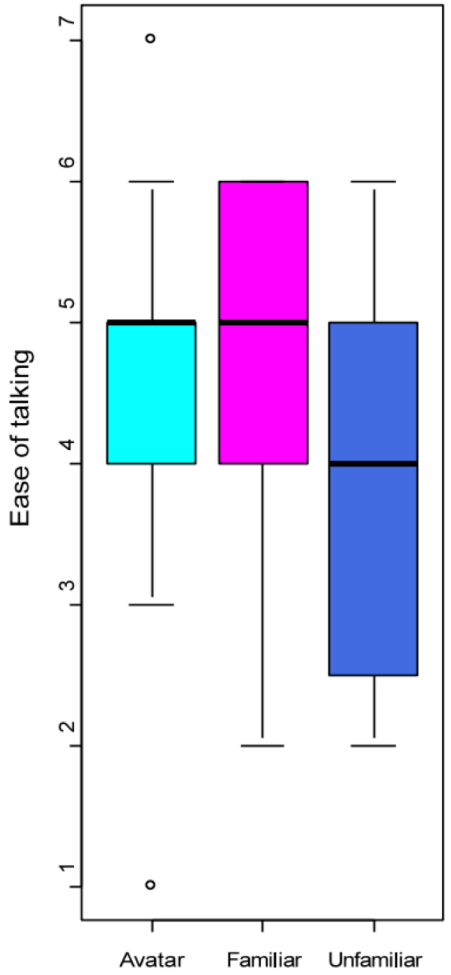


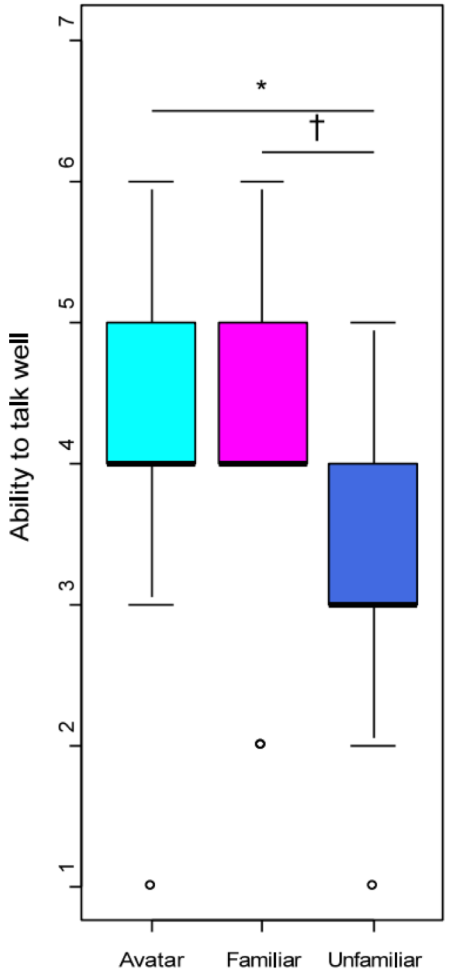


Human-human and human-agent interaction



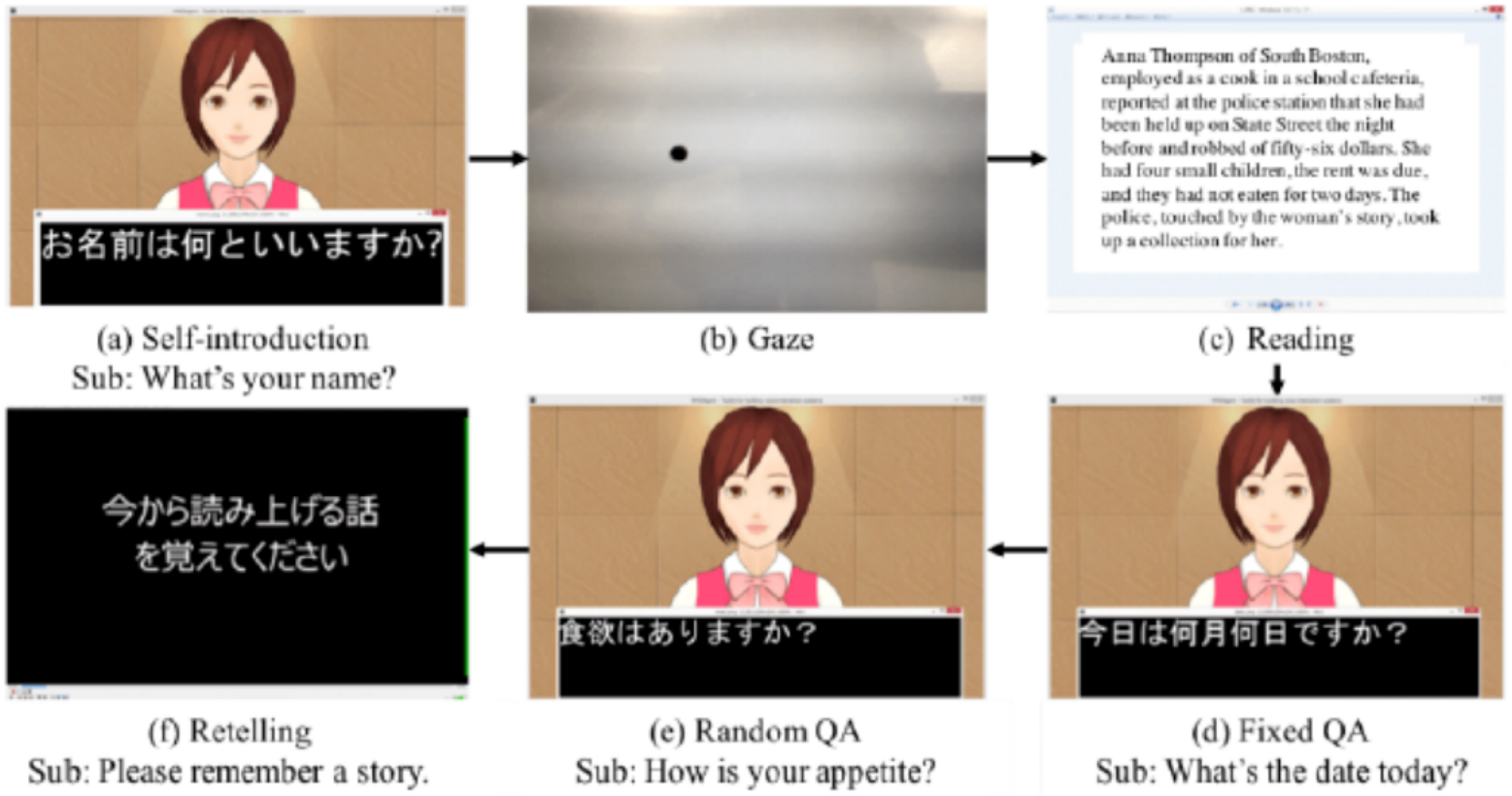






Detection of early stage of dementia

10 minutes interaction: 12 early stage of dementia and healthy control



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

(b) Gaze

(c) Reading

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

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

(f) Retelling
Sub: Please remember a story.

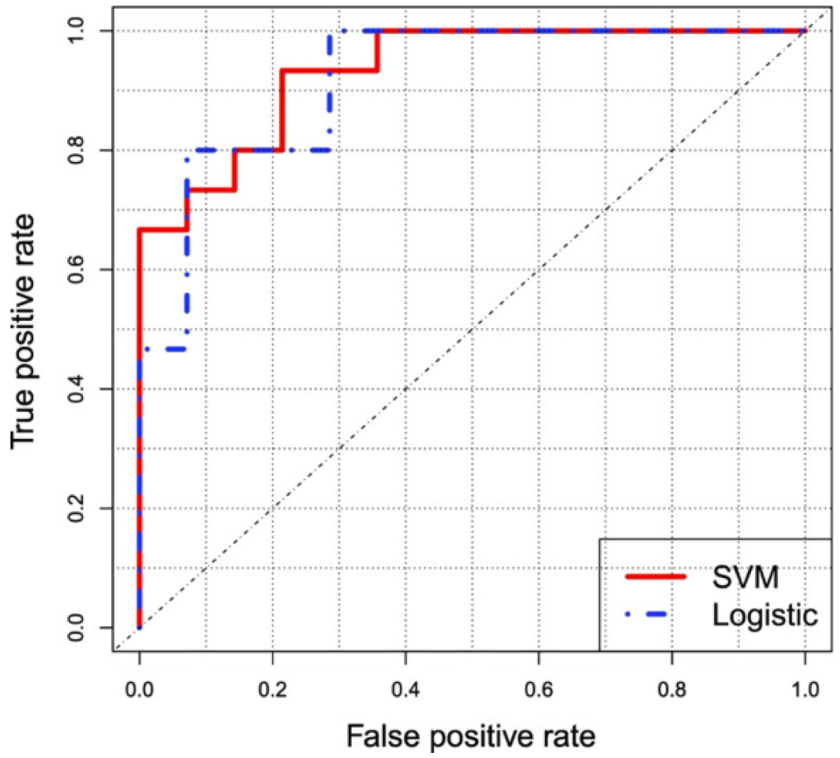
[Tanaka et al., 2017]

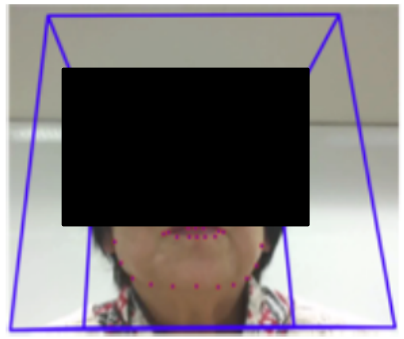
Multimodal-based classifiers

Answers to fixed questions: e.g. what is the date today?

Areas under the ROV curve

- 0.90 (SVM), 0.88 (Logistic regression)





Features: response gap, language, speech, face etc

TAPAS: Goal and contribution

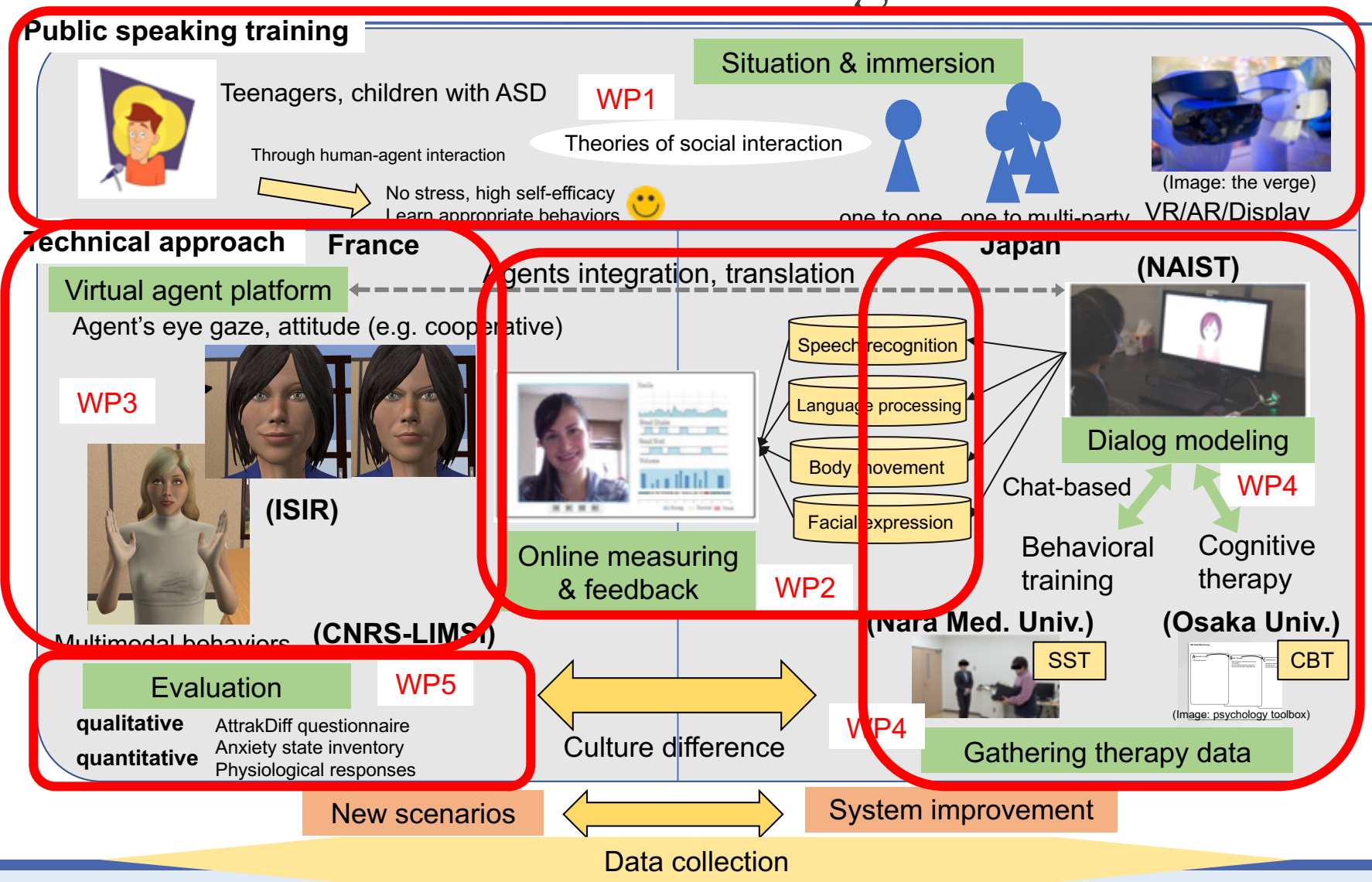
Verbal / non-verbal interactive training system by ECA in everyday situation

- Theoretical framework of social interaction inspired by social psychology: **WP1**
- Behavioral emotion sensing, behavior measurement: **WP2**
- Virtual agent that acts as a coach or member of the audience: **WP3**
- Collect and model training and build dialogue based on SST and CBT: **WP4**
- Scenario setting and evaluation: **WP5**

Expected Contributions

- Evidence-based SST and CBT in VR environment by non-verbal / verbal integrated dialogue
  - Dialogue system research with the above functions has not been developed
- Online / summary feedback, dynamic training according to user's skill
  - Considering verbal / non-verbal information, SST- and CBT-based online and summary feedback are new.
- Communication training with various difficulty levels
  - There is no research corresponding to one-to-one (coach / audience role), one-to-many, many-to-many situations, and difficulty levels
- Comparative consideration of cultural differences between Japan and France

Organization



Nakamura G NAIST (WP2, WP3, WP4, WP5)
Satoshi Nakamura, Hiroki Tanaka, Seitaro Shinagawa
Hirokazu Kato, Yuichiro Fujimoto

Iwasaka G Nara Med. Univ. (WP1, WP4, WP5)
Hidemi Iwasaka, Yasuhiro Matsuda, Kosuke Okazaki

Kudo G Osaka Univ. (WP1, WP4, WP5)
Takashi Kudo, Hiroyoshi Adachi
Yukako Sakagami

Pelachaud G (France CNRS-ISIR) (WP1, WP3, WP5)
Catherine Pelachaud
Donatella Simonetti

MARTIN G (France CNRS-LIMSI) (WP1, WP2, WP5)
Jean-Claude MARTIN
Laurence BOLOT
Céline CLAVEL

WP1: Theoretical framework

Collaboration with other research groups

SST scenario
- Bellack model
- Self-introduction, motivation, modeling, role-playing, feedback, homework

CBT scenario
- Beck model
- Cognitive reconstruction

Based on the above scenarios

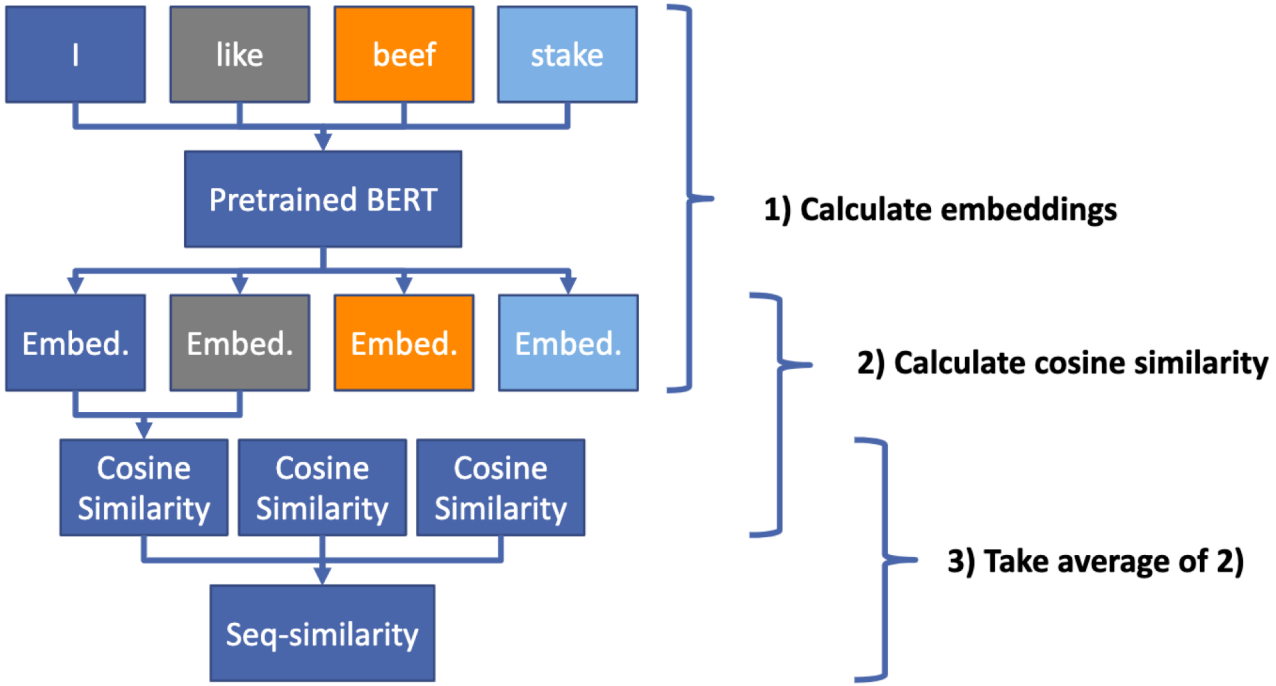
Virtual agent adaptively change behaviors as human do

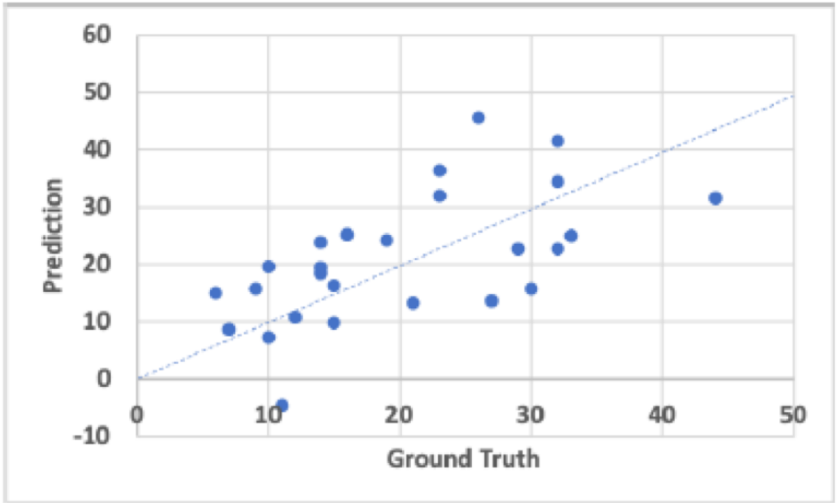
WP2: Behavior sensing on SST

Predicting Social Responsiveness Scale-2 (SRS2)

Input BERT similarity, speech and action units（n=27） [Saga et al., 2020]

Correlation coefficient for SRS2 = 0.6





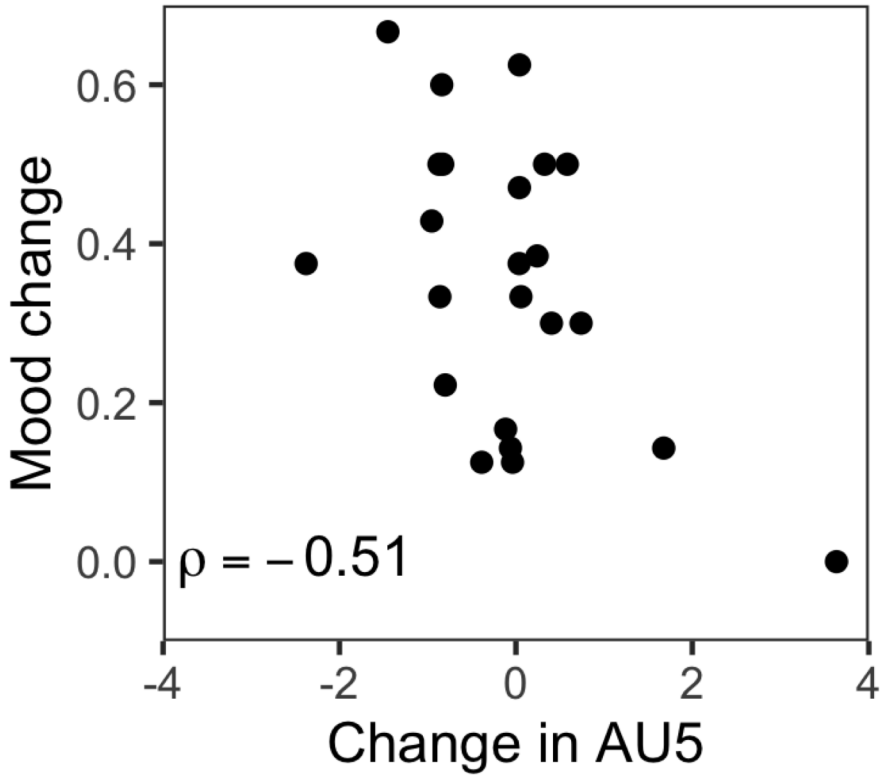
Feature weights

Name	Coefficients
Fluency	-595.7
Energy	74.0
F2/F1 Mean	32.5
F2/F1 SD	-30.2
Int SD	13.6

WP2: Behavior sensing on CBT

Analysis of action units during CBT with virtual agent

AU5 is related to the mood change (n=23) [Shidara et al., 2020]



WP3: Virtual agent

Platform: Greta Ogre3D

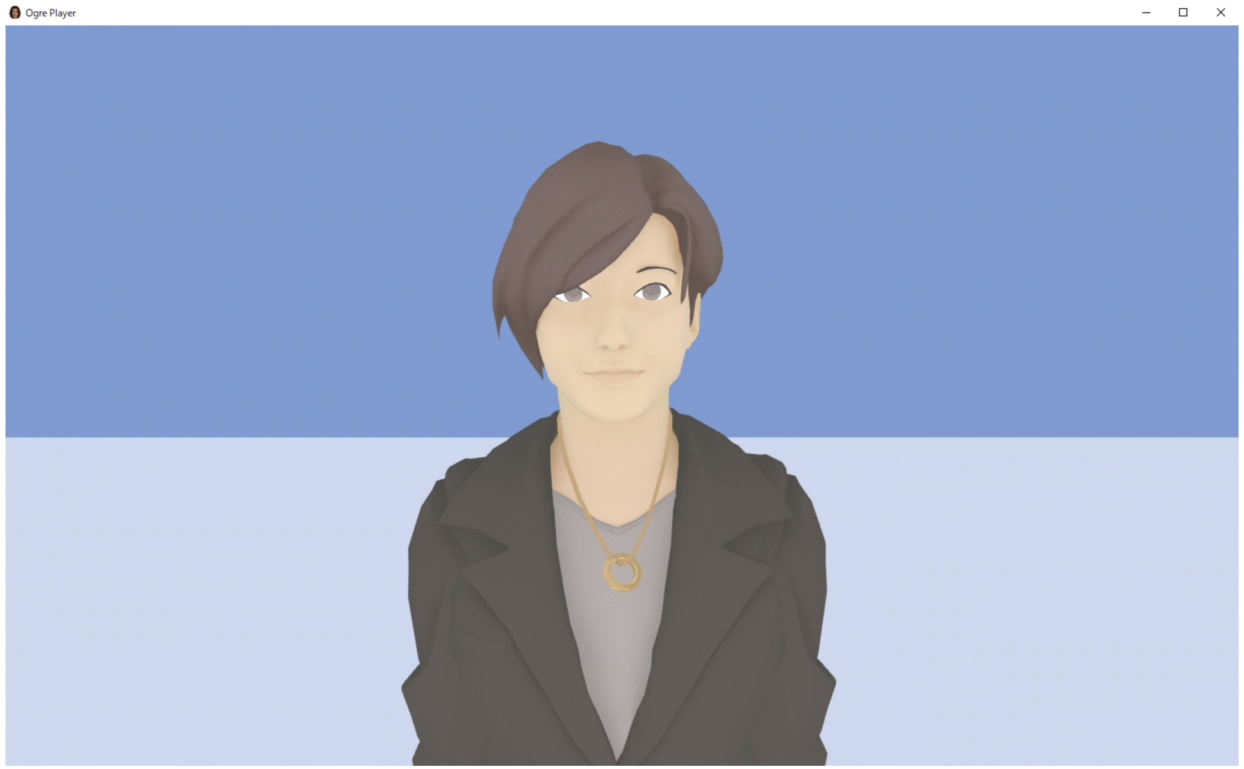
Speech recognition: Google ASR

Speech synthesis: CereProc TTS、Lip synch

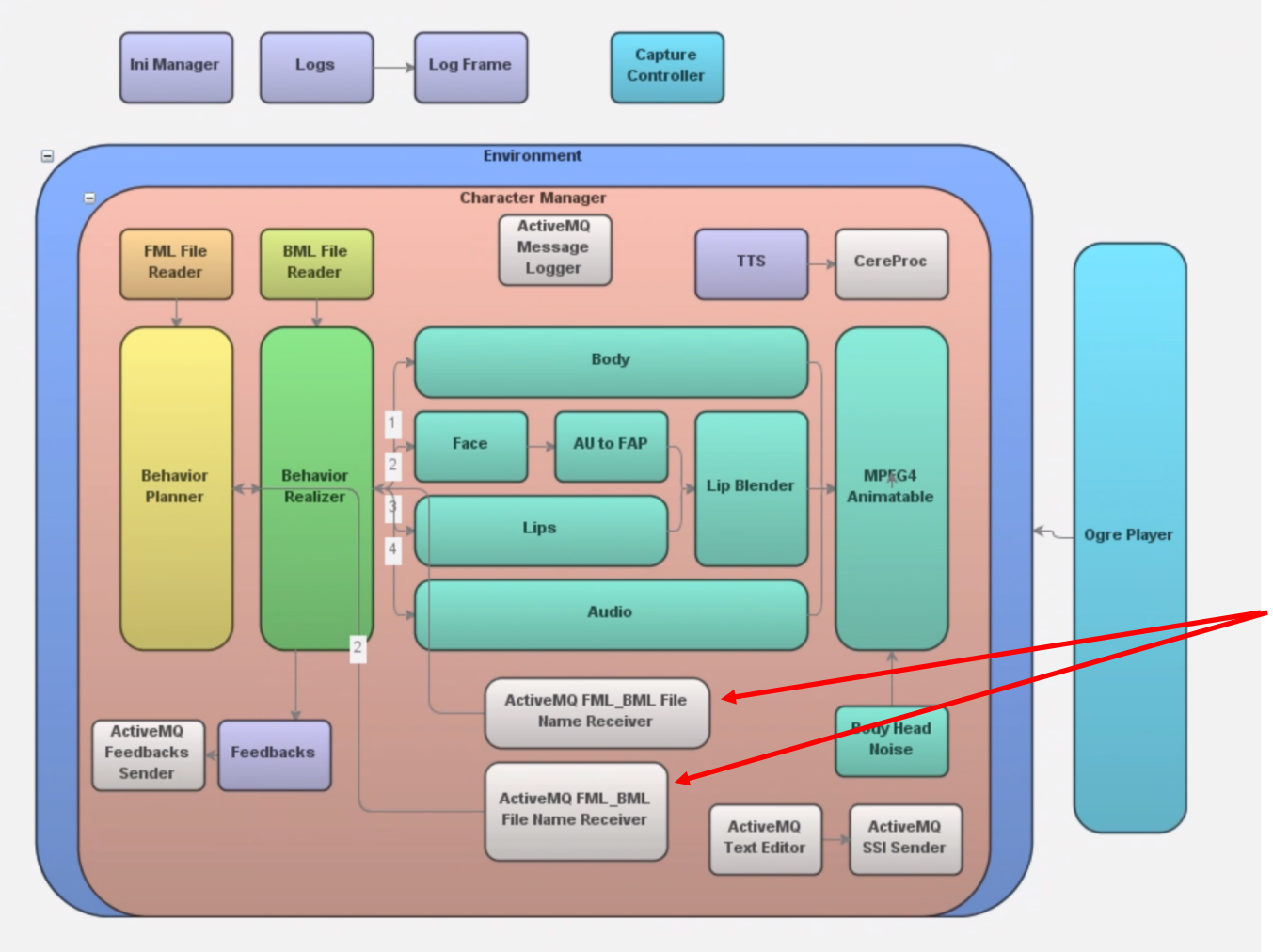
Connection: Java ActiveMQ

Character: Original（anime and real）

Head nodding



Greta details



Thrift: connect to Greta Unity 3D
Greta Unity 3D

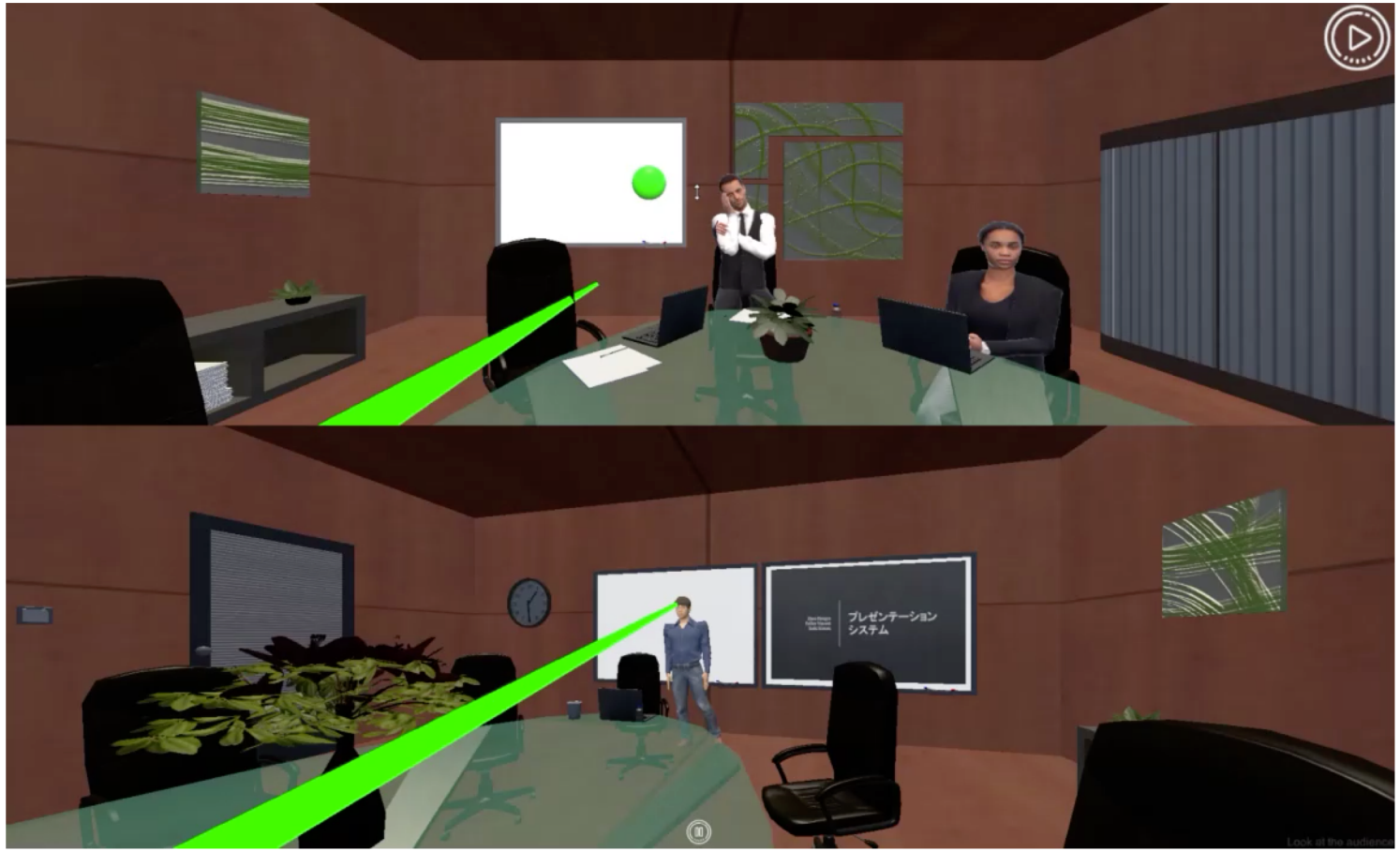
- Lower body movement
- Implement to Hololens

Speech recognition

Connection

Greta Unity towards VR/AR

Presentation in head mount display [Zhou et al., 2020]



WP4: Social skills training（SST）

Medically established method for behavior training

Flow: Situation setting-> Motivation -> Modeling-> Role play-> Feedback-> Homework

1 period: about 1 hour × 10 times

Language: words, expression, consistency, coherence, causality, etc.

Speech: pitch, speaking speed, filler, voice quality, amplitude etc.

Image: facial expression, gesture, posture, eye gaze etc.

SST data collection

Start 2020.9-、annotation, diagnosis (ASD, Schizophrenia)

Due to COVID-19, need partition

2*Kinect
2*Face
1*whole

Dialogue

Nara Medical Univ. and NAIST

Hello. I'm tapas. What is your name?
Hello, hundred's. How have you been recently?
Was that so, today we will learn speaking skills together. Why do you think speaking skills are important?
Thank you very much. When you have a difficult time, you tend to focus on yourself. You can increase your sense of solidarity with others by looking at others and the good things and communicating them appropriately.
...
Let's do role-playing.
Tell me your recent fun story.
Then please.
Yes, this is all right. Thank you very much.
We will calculate the result, so please wait.
Now for feedback.
...
Use the skills you learned today in your daily life.
Thank you very much. See you again.

Cognitive behavioral therapy（CBT）

CBT: Cognitive Behavior Therapy

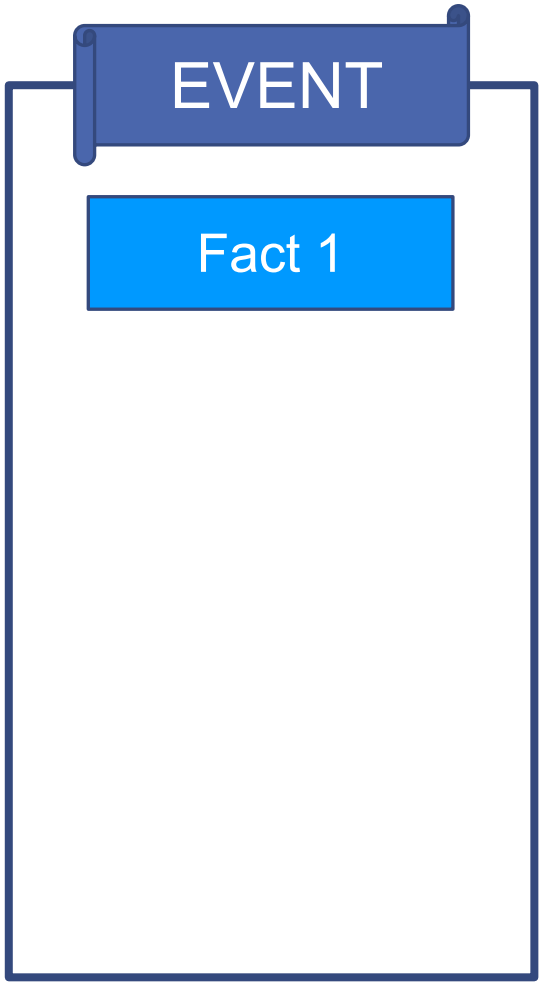
- Medical method to change the schema (cognitive tendency, belief)
- Psychotherapy-based instructional method
- Cognitive Reconstruction, Behavioral Activation, Situation Analysis, Problem Solving, Progressive Relaxation, Assertion, Schema Modification

This study uses Cognitive reconstruction

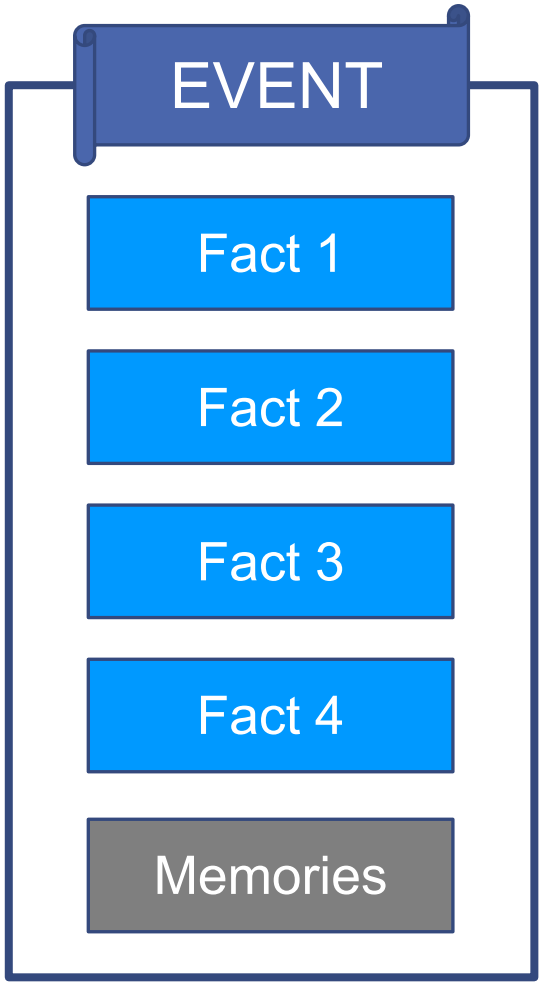
- Change schema depended on only minimal information
- Situation setting-> initial mood-> automatic thinking-> disapproval-> adaptive thinking-> change in mood-> analysis (homework)
- 1 period: 30 minutes or more (usually 4-50 minutes) x 16 times

	Cognitive Therapy Record	ECA Dialogue
Situation	Pointed out my small voice in presentation, and weakness logicality by the audience.	Situation breakdown
Initial score	Discouraged（80）、Sad（90）	Describe emotion state
Negative thinking	I am not able to present in front of audience. I am not good for this job.	Extract tendency
Evidence for this thinking	My voice is small in general. I am not able to explain logically.	Extract evidence
Alternative thinking	There are many people who give a good presentation in small voice. It is possible to present logically if I prepare well in advance.	Provide information to guide for extracting counter thinking.
Change thoughts	I can give a good presentation even voice is small. I will prepare more in advance	Extract alternative thinking
score	Discouraged（20）、Motivated（80）	Recognize change of thoughts

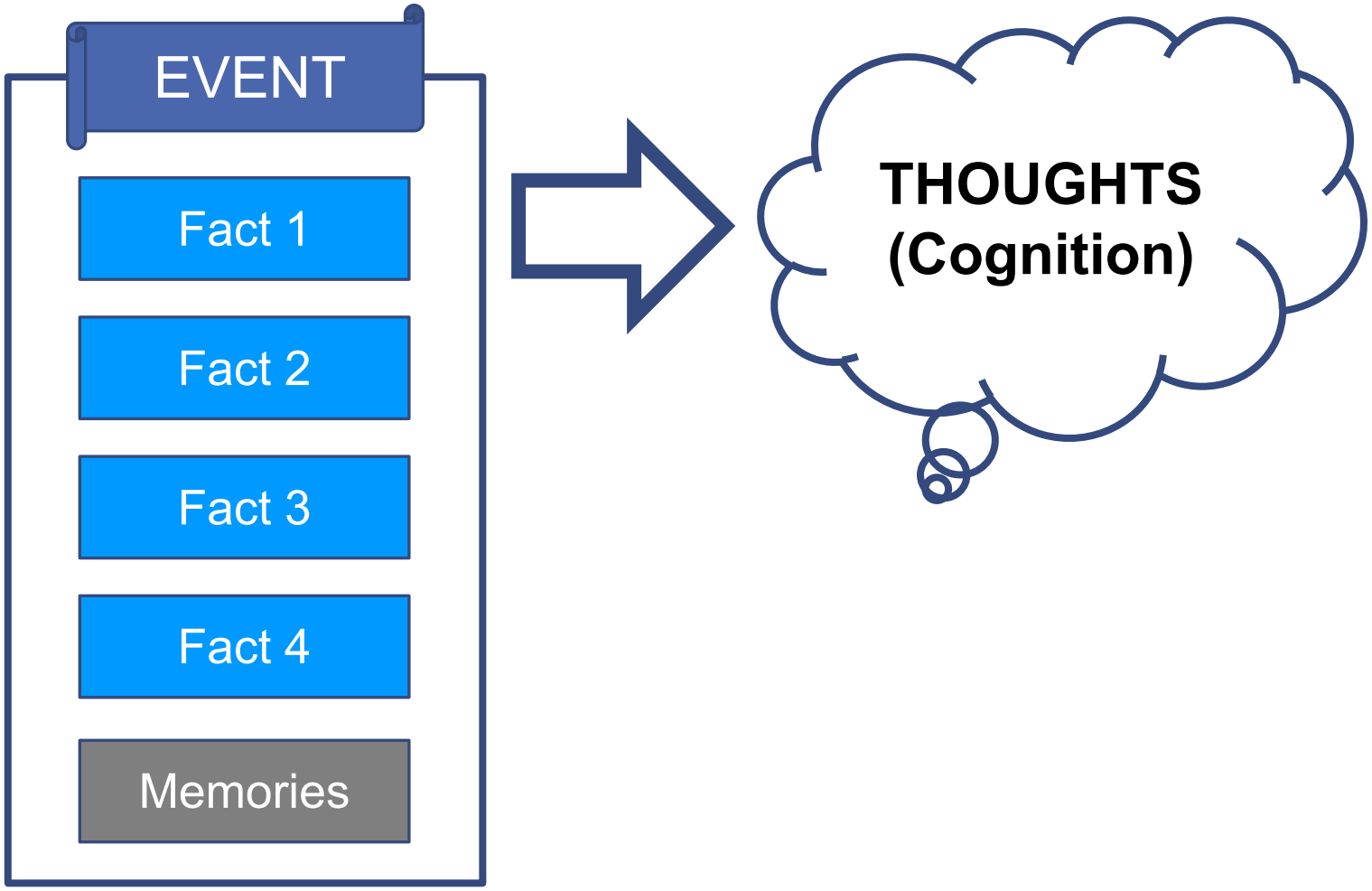
Event and Fact



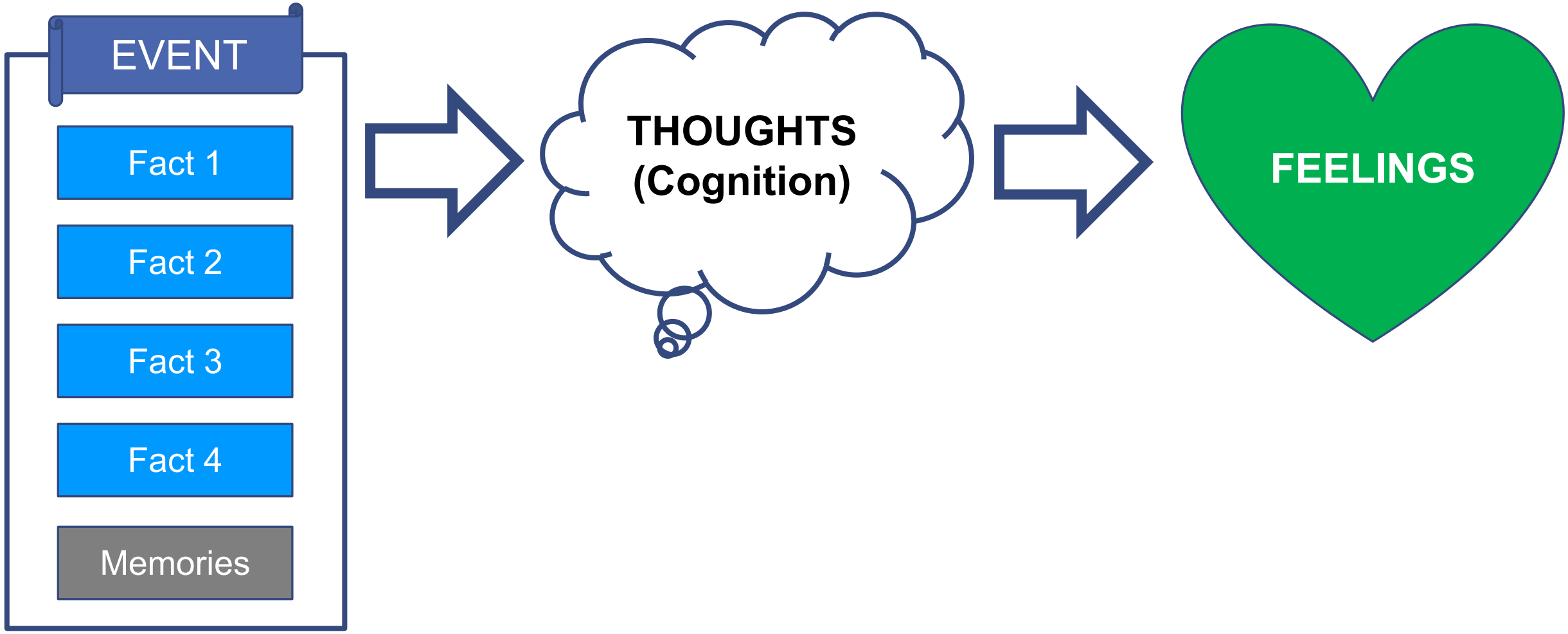
Facts and memory



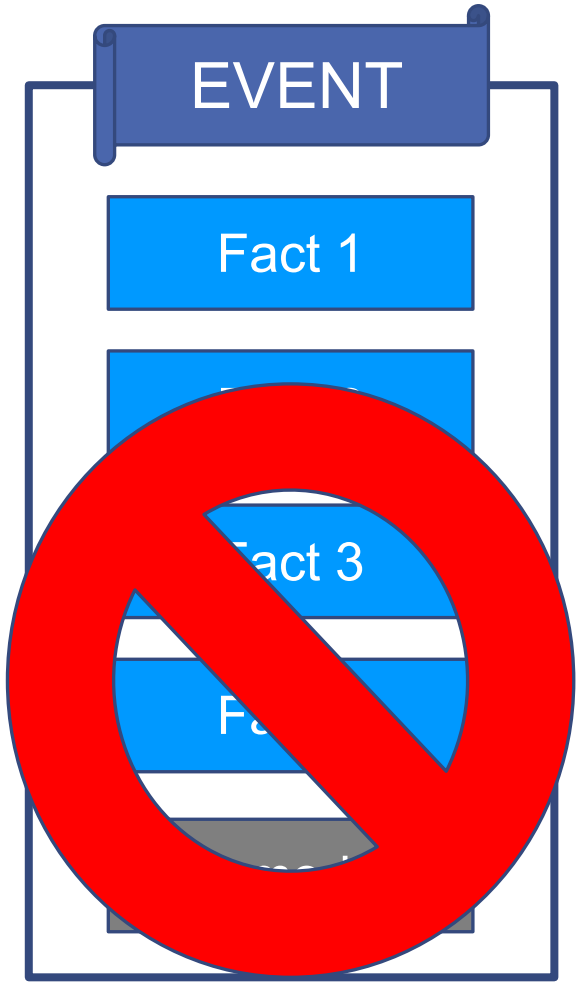
Thoughts


EVENT
Fact 1
Fact 2
Fact 3
Fact 4
Memories
THOUGHTS
(Cognition)

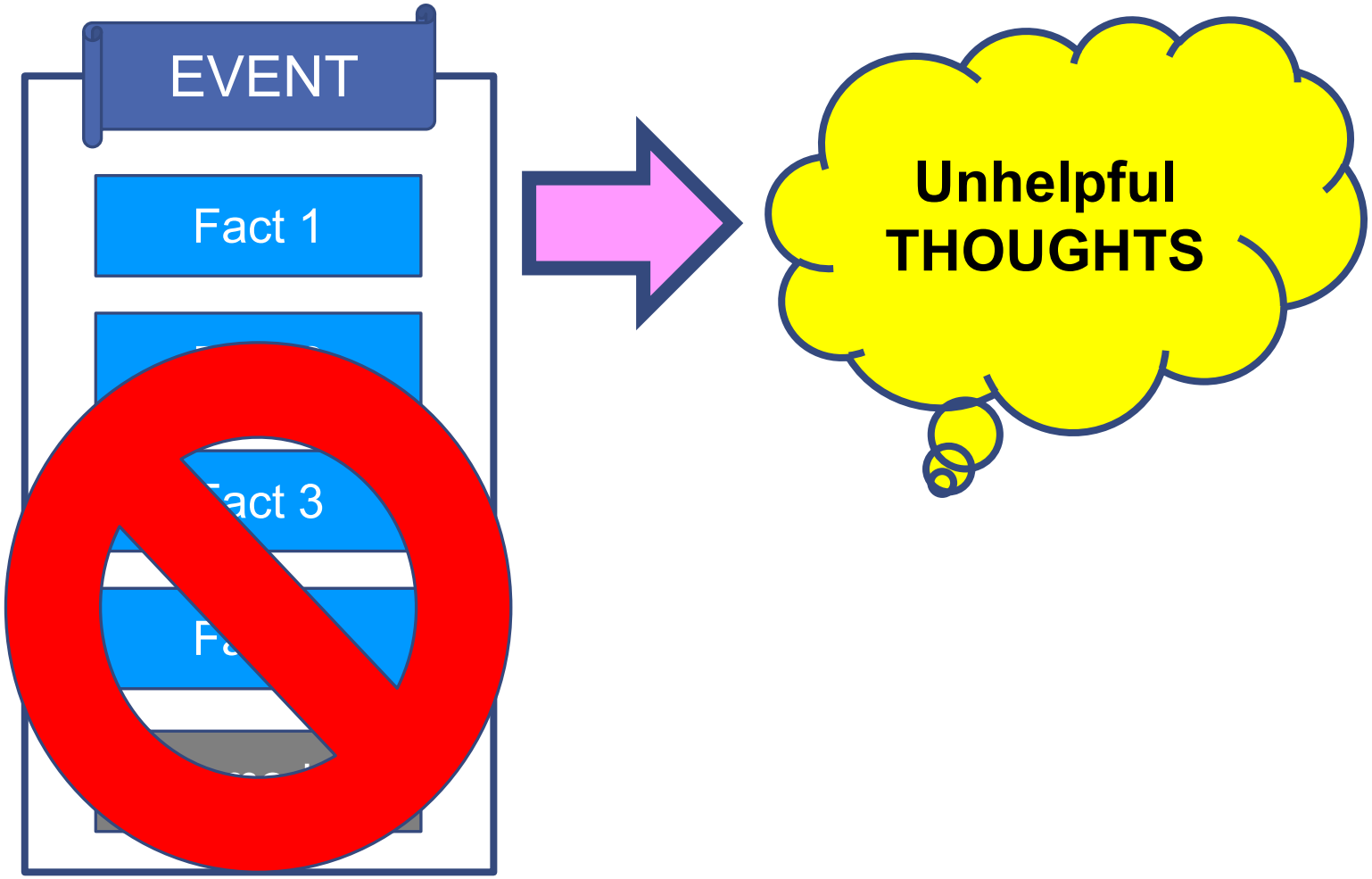
Feelings


EVENT
Fact 1
Fact 2
Fact 3
Fact 4
Memories
THOUGHTS
(Cognition)
FEELINGS

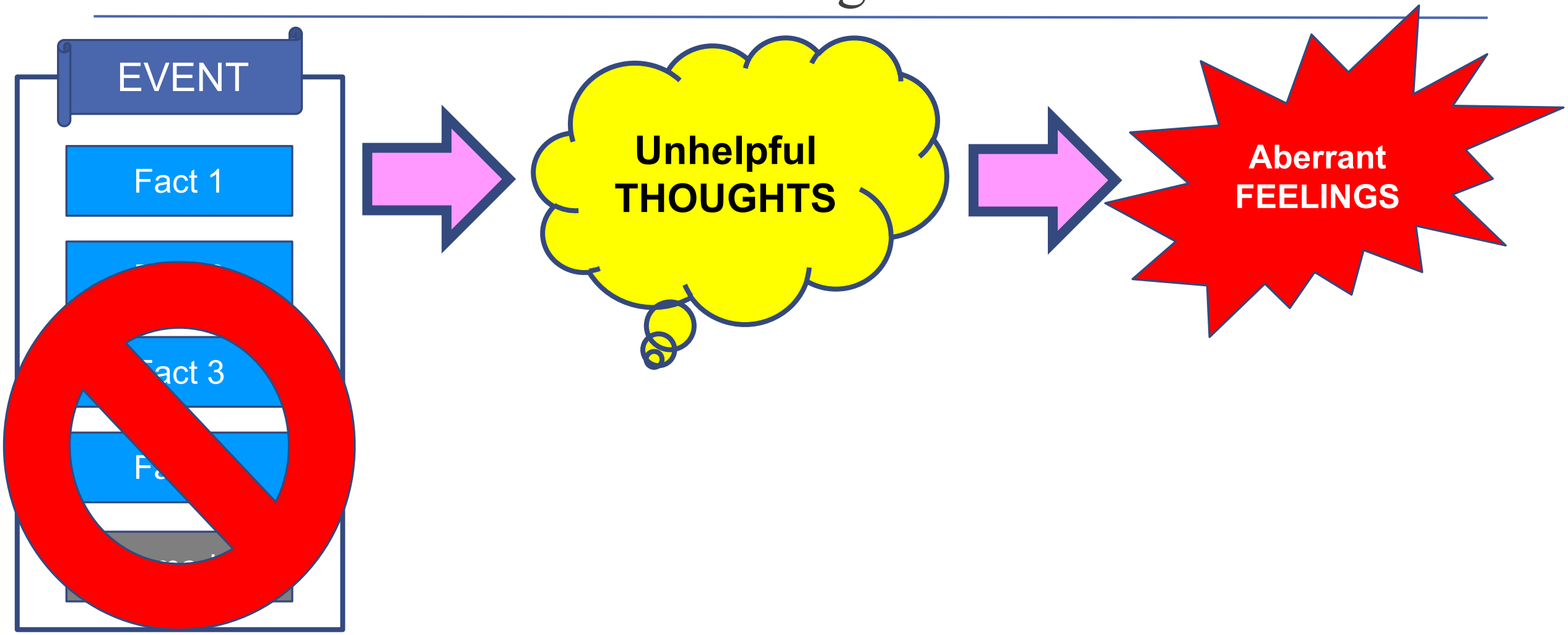
Limited facts



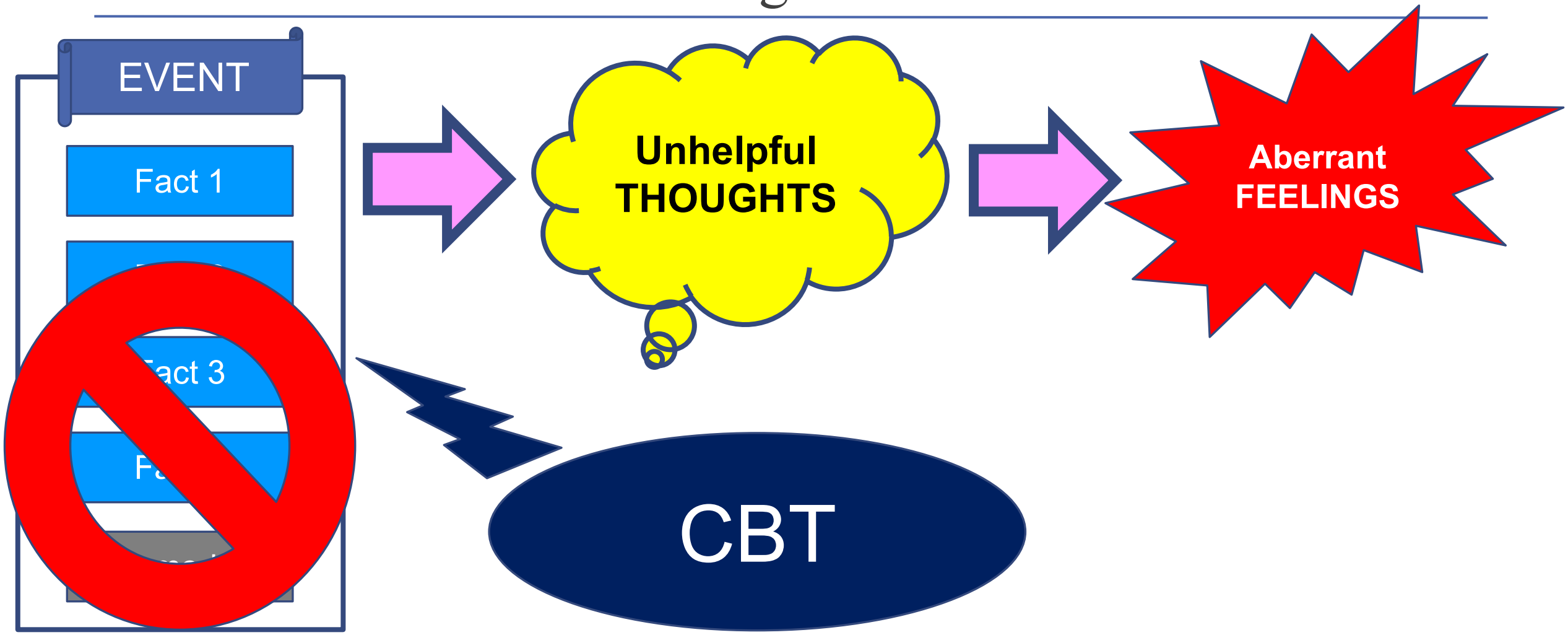
Unhelpful thoughts



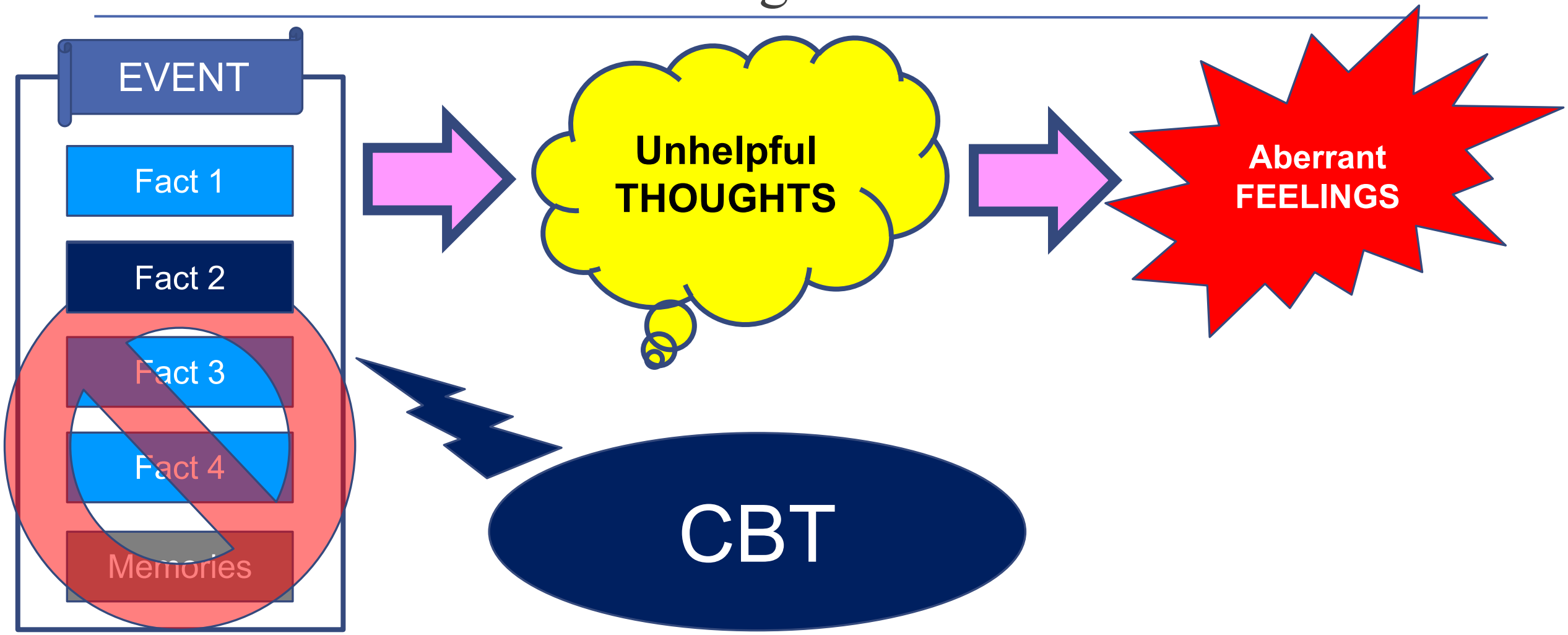
Feelings



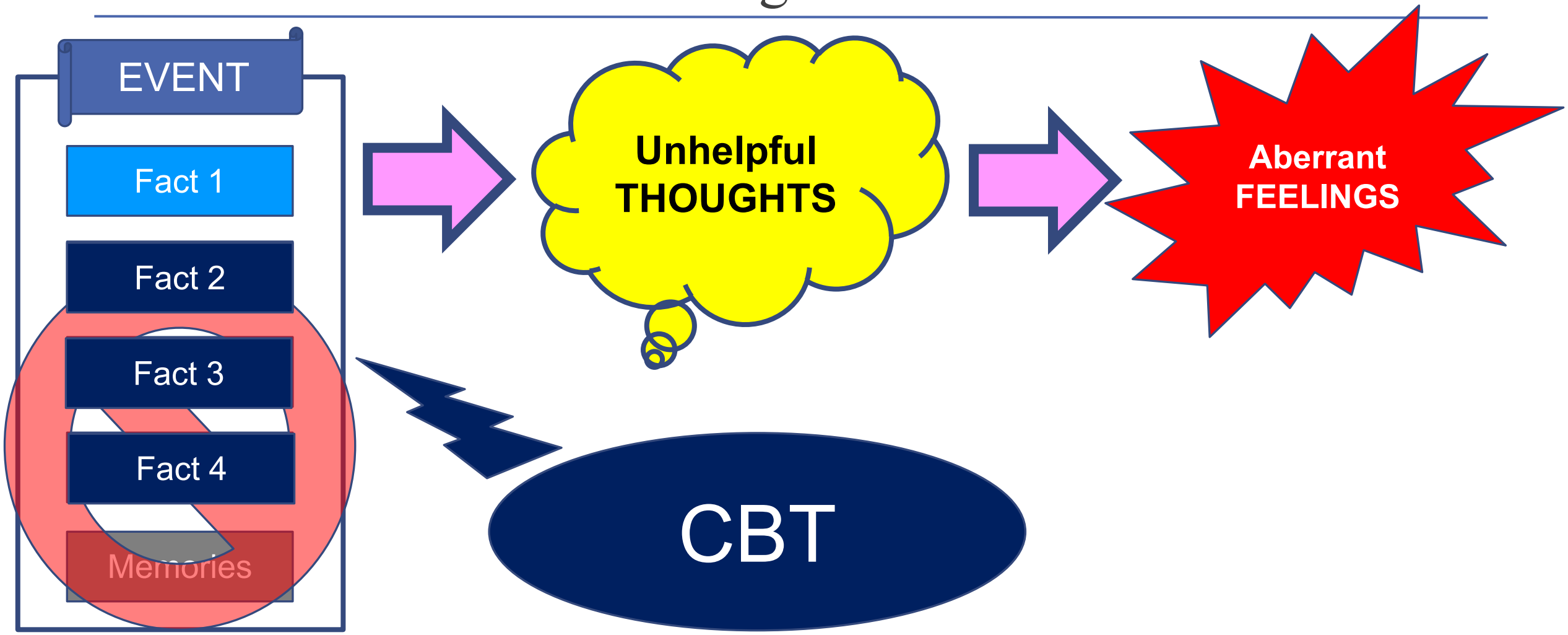
Providing new facts



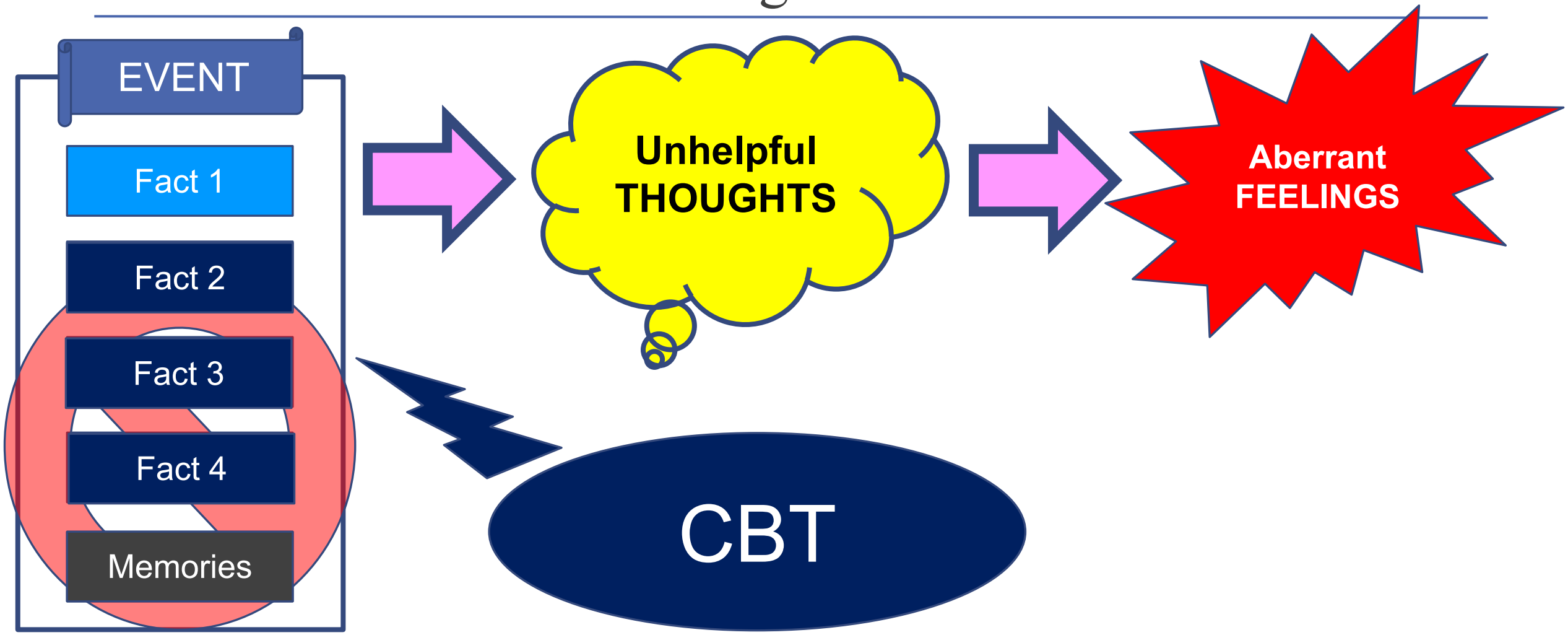
Providing new facts



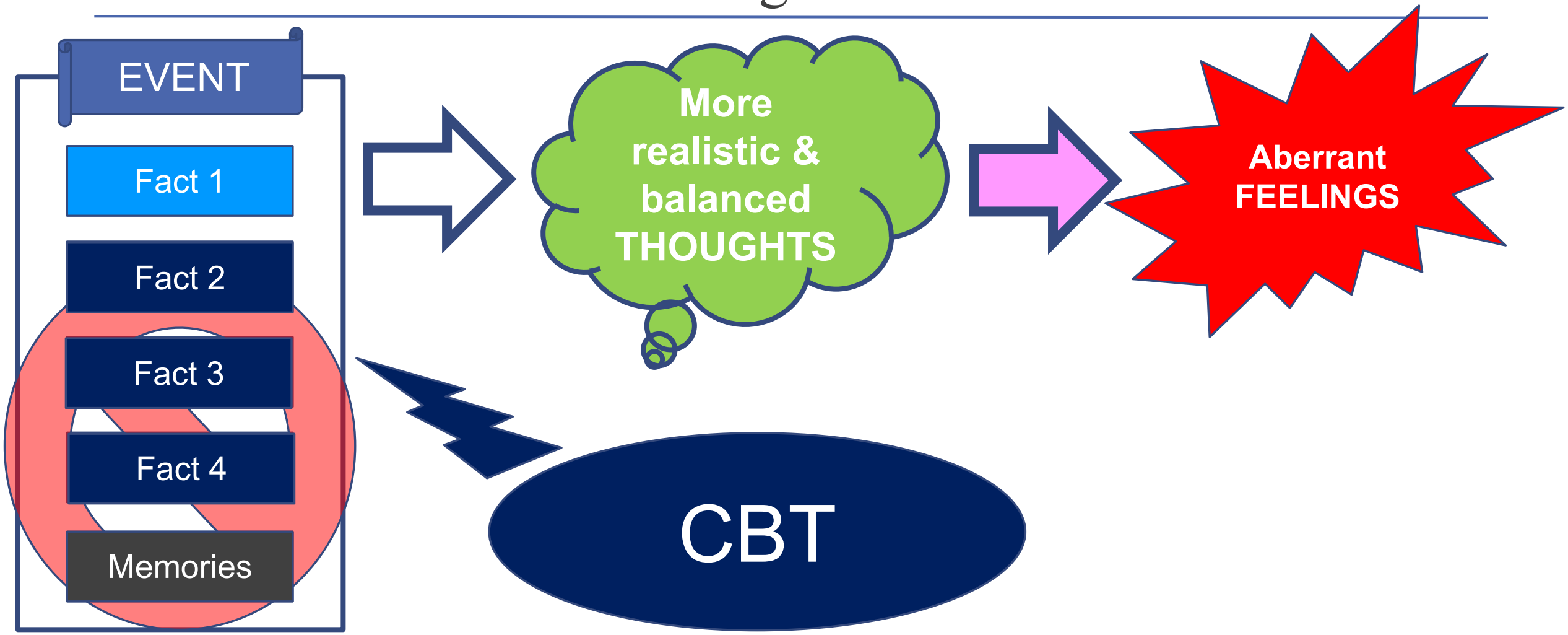
Providing new facts



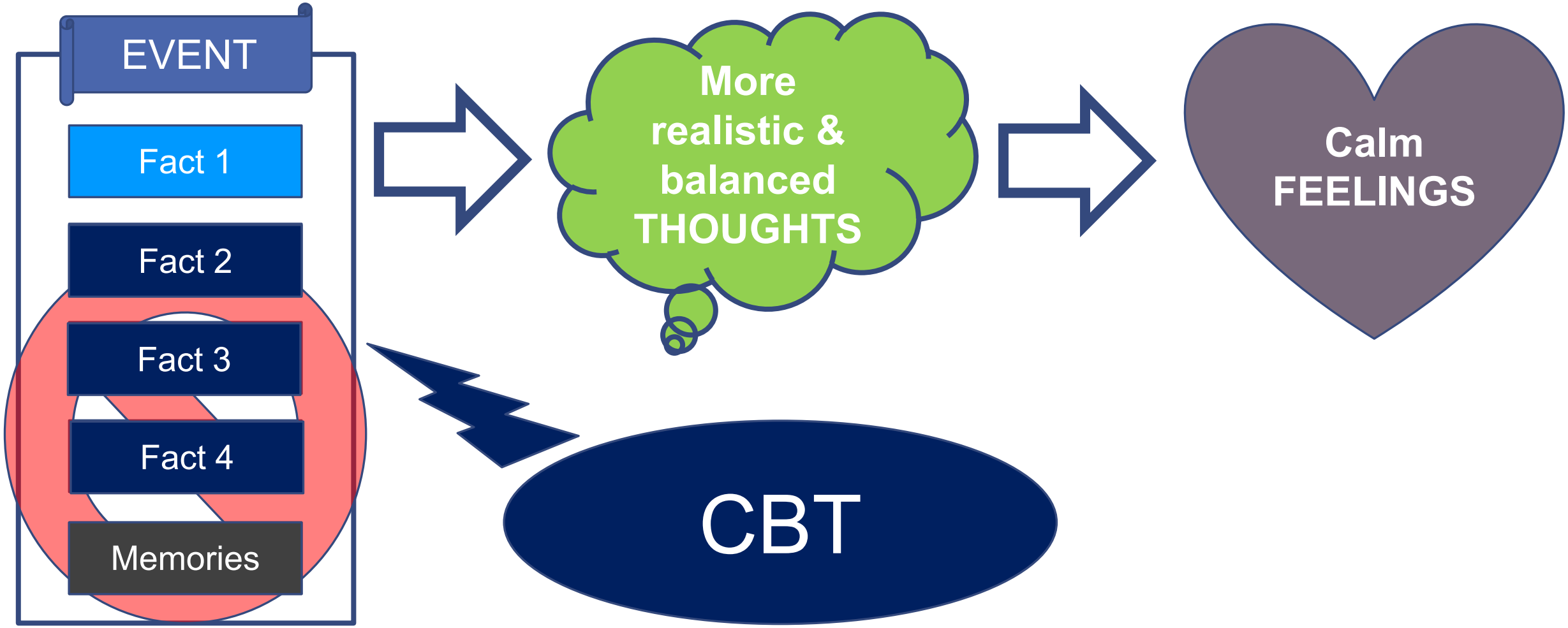
Providing new facts



Providing new facts



Providing new facts



Basic dialogue

Osaka Univ. and NAIST

Item	Sentence
Situation	Hello, my name is Mai, and I am a therapist. Come join me in training to face your worries. Is there anything that's been bothering you lately that's been difficult to deal with or to face? Please tell me something that you think is painful or burdensome.
Mood	How did your mood change at that time?
Mood score (0 to 100%)	How would you describe the intensity of your mood from 0 to 100?
Automatic thought	I want to know why you felt like that in that situation. What thoughts came to your mind when you faced that event?
Proof	I see... Such thoughts that surface when faced with certain events are called automatic thoughts. Since automatic thoughts are unconscious, they sometimes seem true. If your automatic thought is correct, what do you think it is based on?
Disproof	So, do you have another view about that situation? For example, does anything contradict your automatic thinking?

.
.
.

CBT experiment

Virtual agent（MMDAgent [Lee, et. Al., 2013]）

- Multimodal interaction (+text-chat)

Participants

- 23 university students

Interaction with virtual agent

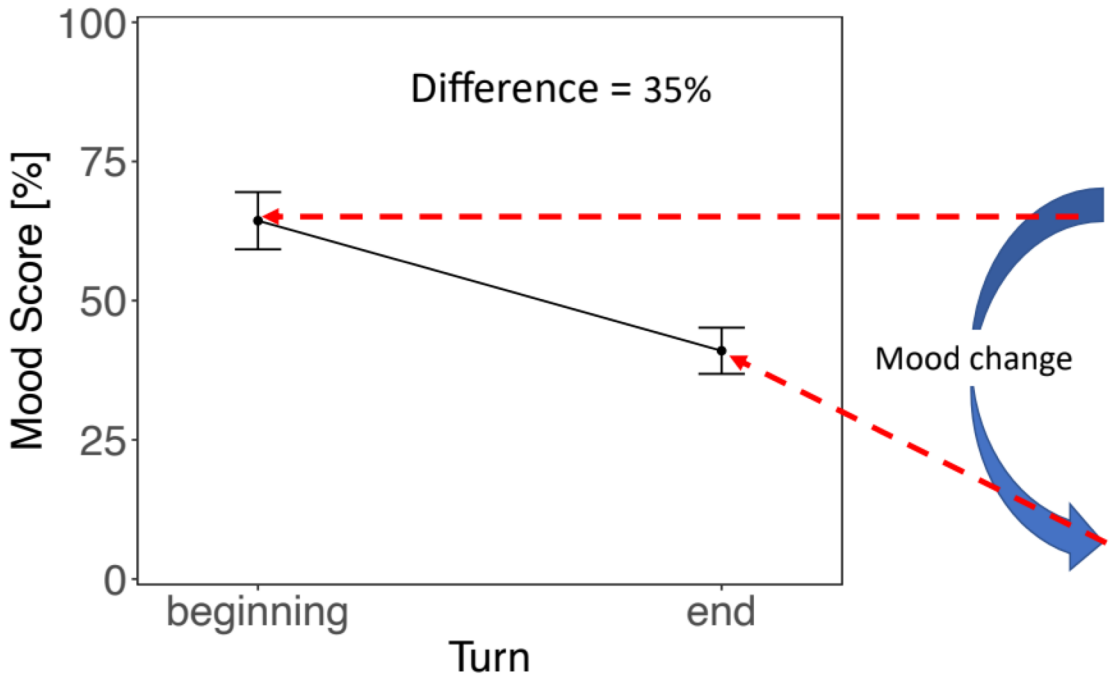
- Same procedure dialogue
- Turn change is done by pressing buttons

Mood score improvement

Human-based CBT [Persons et al., 1980] used mood change

Mood change during CBT by the virtual agent*

*Decrease in negative feelings (sad, etc)



Item of question	Answer example
Situation [What happend?]	I don't get an email reply from a friend.
Mood(mood score(%)) [How do you feel?]	Sad （80%）
Automatic thought [What thoughts came to mind?]	I was hated.
Proof [What is the basis to support the fact?]	No reply since yesterday.
Disproof [Do you have the opposite fact?]	He has been busy these days.
Balanced thought [What do you think flexibly?]	He might have missed the incoming mail.
Mood change (%) [Have you changed your mood?]	Sad （30%）

Evaluation

Assessment

- Social Responsiveness Scale（SRS2）
- K6
- QIDS
- Toronto
- Self efficacy
- ADOS
- Emotion Discrimination task etc

EEG: detecting incongruity and affective state

Eye tracking study

ICMI workshop (SAMIH)

https://sites.google.com/view/wsamih/home

Reference

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.

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.

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.

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.

Takeshi Saga, Hiroki Tanaka, Hidemi Iwasaka, Satoshi Nakamura, Objective Prediction of Social Skills Level for Automated Social Skills Training Using Audio and Text Information, Adjunct of the 2020 International Conference on Multimodal Interaction (ICMI), Workshop on Social Affective Multimodal Interaction for Health, Oct. 2020.

Kazuhiro Shidara, Hiroki Tanaka, Hiroyoshi Adachi, Daisuke Kanayama, Yukako Sakagami, Takashi Kudo, Satoshi Nakamura, Analysis of Mood Changes and Facial Expressions during Cognitive Behavior Therapy through a Virtual Agent, Adjunct of the 2020 International Conference on Multimodal Interaction (ICMI), Workshop on Social Affective Multimodal Interaction for Health, Oct. 2020.

CBT data [Beck and Beck, DVD]