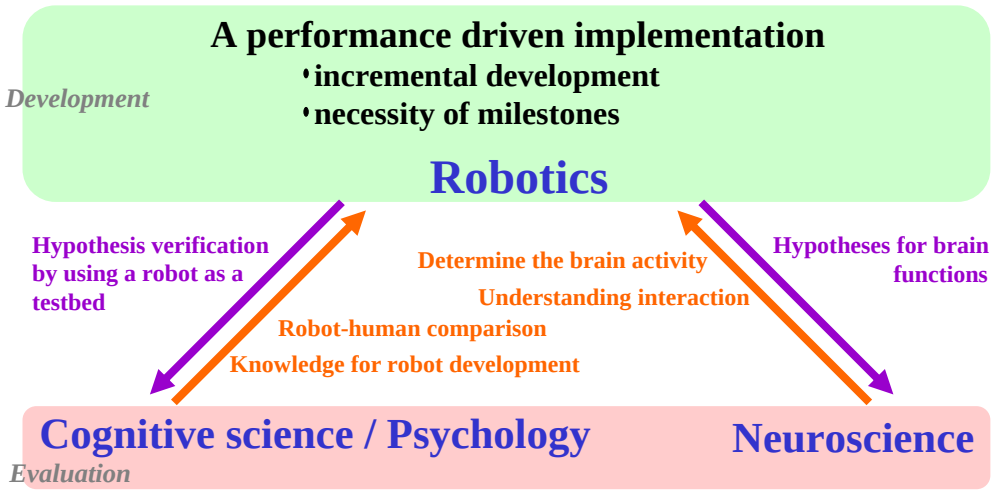
Evaluation of the interactive robots

Interdisciplinary research approach to developing interactive humanoid robots



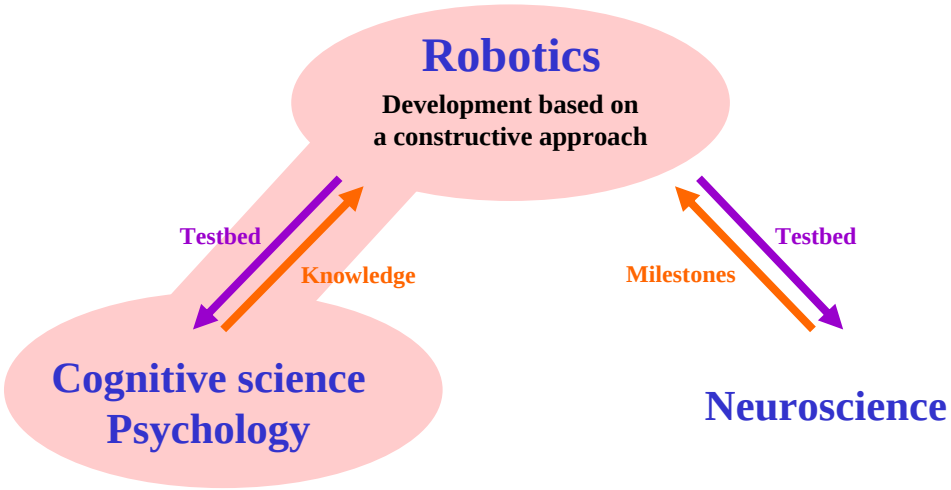
How can we evaluate the interactive robot?

1. Multiple incomplete tasks
2. Importance of subjective impression
3. How long can we keep interacting

➡ New evaluation criteria

Evaluation

Short-term experiments for subjective and objective evaluations

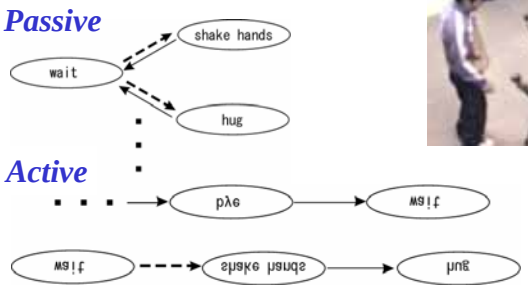


Psychological Evaluation

Passive:
The robot waits until a subject interacts. When the subject touches the robot, it exhibits one of the friendly behaviors. Then it waits again.

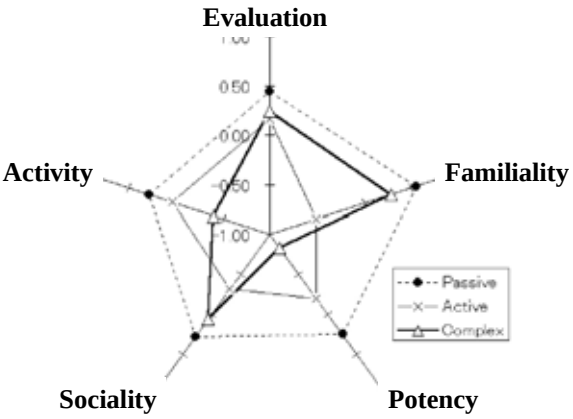
Active:
The robot asks interaction to a subject and continues the friendly behaviors while the subject reacts to the behaviors.

Complex:
In addition to the *Active pattern*, it sometimes exhibits *Idling* and *Daily work* (move around) behaviors instead of waiting.



Psychological Evaluation: Factor Analysis

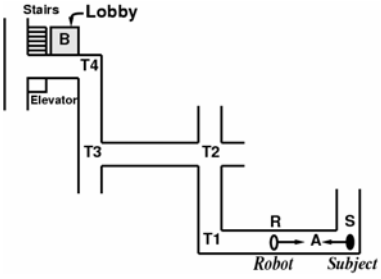
	Evaluation	Familiarity	Potency	Sociability	Activity	Complexity
Good	**0.808**	0.252	0.127	0.003	0.146	0.755
Kind	**0.759**	0.110	-0.038	0.197	-0.209	0.671
Pretty	**0.690**	**0.514**	-0.111	0.132	-0.020	0.730
Exciting	**0.505**	0.442	0.153	0.182	0.355	0.633
Warm	0.451	0.074	0.110	0.399	0.117	0.394
Accessible	0.417	0.136	0.240	0.246	-0.085	0.318
Humanlike	0.336	0.116	0.167	0.312	0.137	0.271
Pleasant	0.186	**0.873**	0.112	0.166	-0.011	0.837
Friendly	0.313	**0.727**	0.059	0.175	0.001	0.660
Likable	**0.573**	**0.614**	-0.108	0.086	0.149	0.746
Cheerful	0.021	**0.527**	0.029	0.351	0.078	0.408
Favorable	0.453	**0.589**	0.130	0.302	0.002	0.573
Intelligent	0.331	0.448	0.252	-0.263	0.254	0.508
Showy	-0.033	0.079	**0.667**	0.263	0.279	0.599
Complex	0.012	0.021	**0.654**	-0.295	0.113	0.528
Sharp	0.112	-0.004	**0.648**	0.079	0.152	0.451
Full	0.226	0.339	**0.516**	0.203	0.213	0.519
Light	0.239	0.458	0.173	**0.657**	0.162	0.755
Agitated	0.195	0.202	-0.099	**0.556**	0.224	0.447
Frank	0.480	0.329	0.103	**0.528**	-0.129	0.645
Rapid	-0.024	0.022	0.374	0.040	**0.711**	0.648
Quick	-0.044	0.022	0.204	0.124	**0.574**	0.389
Interesting	0.377	0.351	0.084	0.163	0.390	0.452
Proportion	16.355	15.565	8.901	8.792	6.795	

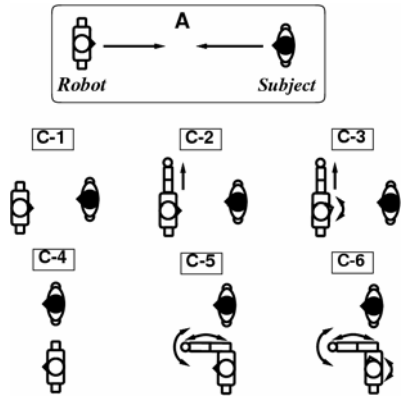


Limitation of SD method …

Mutual Entrained Gestures in Human-Robot Communications

T. Ono, M. Imai, H. Ishiguro, A model of embodied communications with gestures between humans and robots, Proc. 23th Annual Meeting of the Cognitive Science Society, pp. 732-737, 2001.

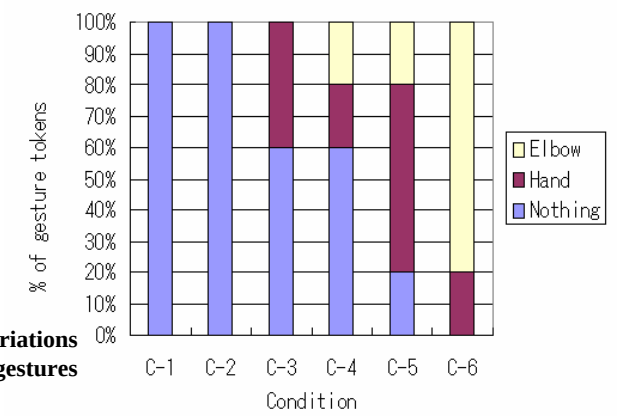




Mutual Entrained Gestures in Human-Robot Communications

Hypothesis 1:
Rich robot behaviors induce various human communicative gestures

Hypothesis 2:
Humans can understand better with the robot gestures



Evaluation by variations of human gestures

	C-1	C-2	C-3	C-4	C-5	C-6
Time to destination	69.5	71.3	67.7	70.2	66.8	65.4
Number of subject not arriving	1	2	2	0	0	0

Limitations of psychological evaluation

Evaluation by SD Method

→ **Difficulties of controlled experimentations using complicated robots**

Evaluation based on behaviors of subjects

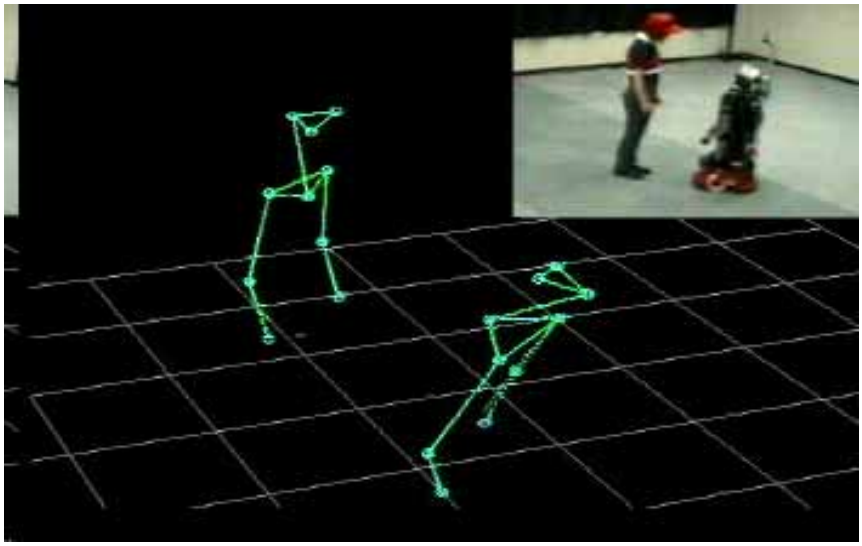
→ **Difficulties for identifying the factors**
Poor human behavior and personal habits

↓

Direct evaluation by measuring brain activities

Precise measurement of Gestures using VICON

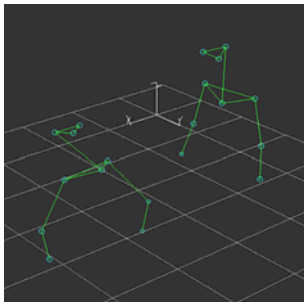
Pattern recognition techniques for psychological study



Experimental setup

- **Settings**
  - **Subject: 26 university students**
  - **Freely interacted with it. 10 min.**
  - **Instruction for interaction**
- **Evaluation**
  - **Body movement**
    - **Motion capturing system**
    - **High-resolution（120Hz, 10mm）**
  - **Subjective evaluation**
    - **Questionnaire**

Attachment of markers



Captured data

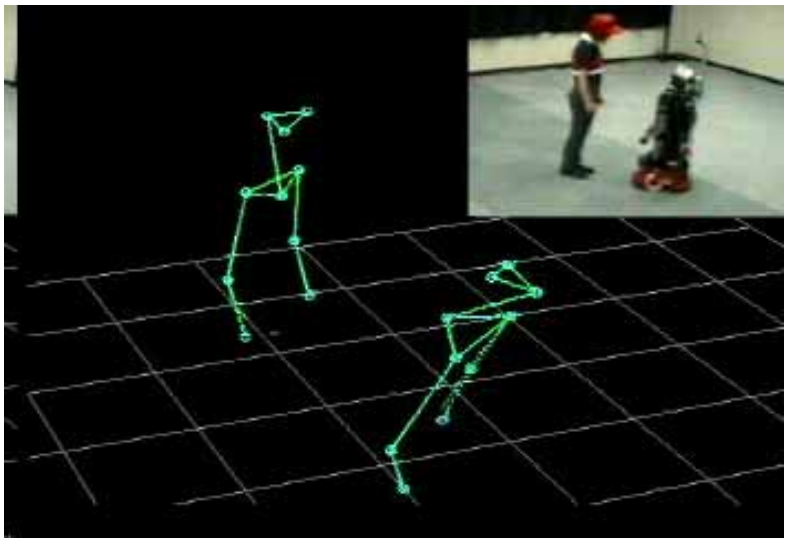
12

Environment

7.5m ✕ 10m room (with 12 cameras)

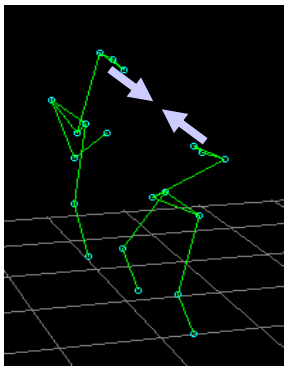
A scene from the experiment

(Data from the motion capturing system)



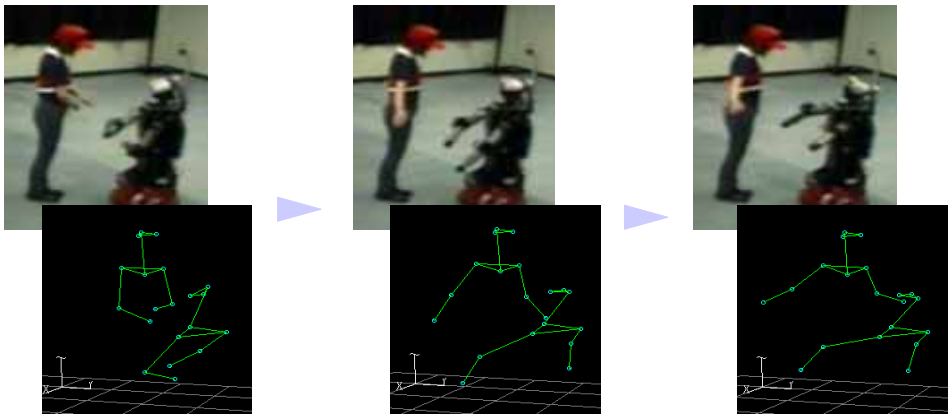
Analysis of body movements

- **Eye contact** **(correspondence of each head orientation)**



Analysis of body movements

- **Synchronized arm movements**
  (correlation of hand position during 3 seconds)



Results for body movements

- **We analyzed each 10 min of captured data**

	Mean	Std. Dev.
Distance (m)	**0.547**	0.103
Eye contact (sec.)	**328**	61.8
Eye height (m)	1.55	0.124
Distance moved (m)	35.2	17.0
Dist. moved by hands (m)	108	29.5
Sync. movements (sec.)	**7.95**	6.58
Touch (num. of times)	54.9	20.8

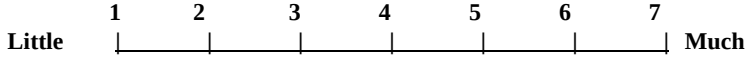
- **Smaller distance than human-human communication (More than 1.2 m during human-human talk)**
- **They made eye-contact more than half of 10 min experiment**
- **Some of subjects imitate robot's arm movement**

Subjective evaluation

- **Using adjective pairs (SD method)**
- **Average of 5 adjective pairs**
  - **Correlation with body movements**

	Mean	Std. Dev.
Good	4.88	0.95
Kind	4.85	1.29
Pretty	5.08	0.93
Exciting	4.46	1.61
Likable	4.77	1.03
score	4.81	0.92

What do you think about this robot?
Please rate each adjective pair on the scale:



Correlation of body movements with subjective evaluation

- **With *Evaluation score*:**
  Eye contact, synchronized movements
- **Among body movements:**
  Synchronized movements-Distance moved by hands (0.61), etc.

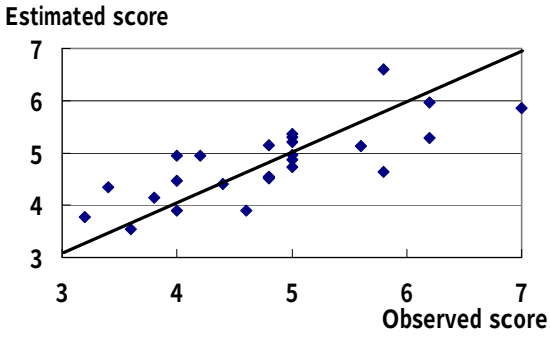
↔**Distance moved by hands – *Evaluation score* (0.01)**

- **It suggests the importance of cooperative behaviors**

Correlation	Eval.
Dist.	-0.04
Eye Contact	**0.57**
Eye height	0.08
Dist. Moved	-0.32
Dist. moved hand	0.01
Synchronized	**0.54**
Touch	0.21

Multiple linear regression analysis

- **Estimation of subjective evaluation with body movements**
- **59% was explained**



	Coefficients
Dist.	0.173
Eye contact	**0.476**
Eye height	0.019
Dist. moved	-0.228
Dist. moved hand	-0.029
Synchronized	**0.535**
Touch	-0.186

Analysis of human-robot interaction with body movement

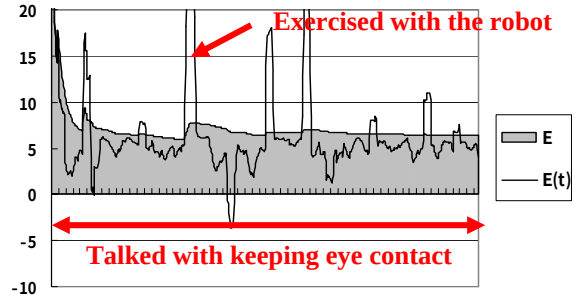
- **Comparison of subjective evaluation and body movements**
  - **We find how body movement correlate with the subjective evaluation**

- **Analysis of Human-Robot interaction with body movement**
  - **It enables *dynamic evaluation***
  - **Even estimation can be useful**
- **Example of Application**
  - **Estimation of momentary evaluation *(dynamic evaluation)***
  - **Evaluation of each behavior of robot**

Momentary evaluation

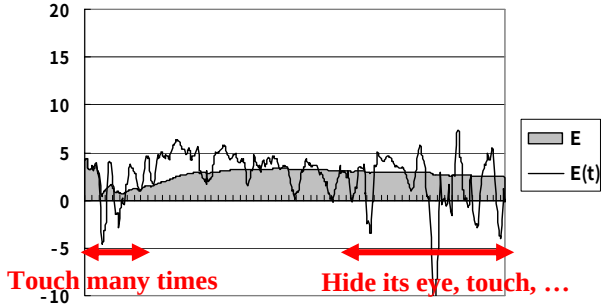
- **Apply the regression to momentary body movement**

$$E(t) = \alpha_{dist} \bullet DIST(t) + \alpha_{ec} \bullet EC(t) + \alpha_{eh} \bullet EH(t) + \alpha_{dm} \bullet DM(t) + \alpha_{dmh} \bullet DMH(t) + \alpha_{sm} \bullet SM(t) + \alpha_{touch} \bullet TOUCH(t) + \alpha_{const}$$



Upper: The subject who mentioned "I nearly regard it as a innocent child."

Lower: The subject who became embarrassed



Evaluation of its behavior

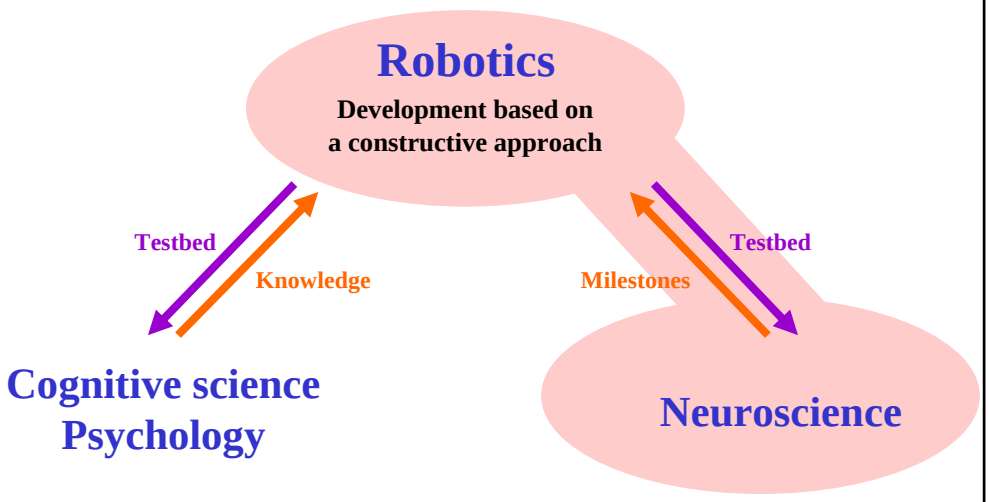
- Average of every subject's momentary evaluation during each module
- Better modules
  - Humans tend to imitate its movement
  - Ask for humans' answer
- Entrain to its interaction
  - Physically
  - Contextually

ID	Contents	Eval.
TICKLE	Tickle	-2.09
APOLOGIZE	Apologize	-1.96
NOT TURN	Say, ``I'm busy,'' and refuse to play together	-0.51
SLEEP POSE	A pose of sleeping	-0.42
STOMACH ACHE	A pose of being fully fed	0.32

Worst 5 situated modules

ID	Contents	Eval.
EXERCISE	Exercise	5.75
ASK SING	Ask humans, "May I sing a sonf?"	5.59
CONDUCTOR	Pose imitating a musical conductor	4.85
WHERE FROM	Ask humans, "Where are you from?"	4.55
LET'S PLAY	Say, "Let's play, touch me"	4.24

Best 5 situated modules



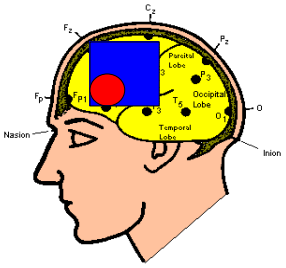
(collaborative works with Prof. Hiraki, Tokyo Univ.)

How dose a human recognize the robot

A primary experimentation by opto-topography

- **Dose the opto-topography measures activities of mirror systems activated by both of action and observation?**
- **Explain the reality of a robot based on activities of the mirror systems.**

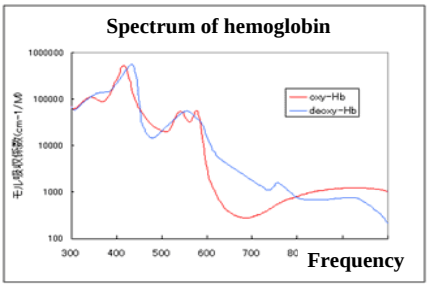
How does the human feel a robot?

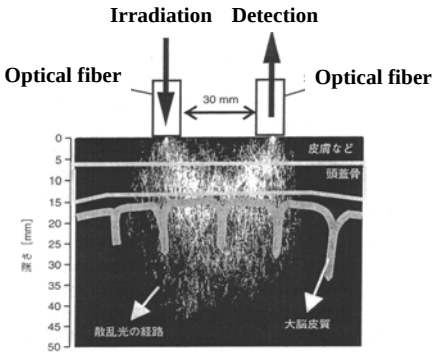


- **We have observed activations in identical areas (motor area) under three conditions, Real human, Real robot and Execution.**
  - ➞ **Mirror systems**
- **The order of active conditions are**
  Execution>Real human>Real robot>Video human>Video robot
- **There is an area that is activated by Real robot and not by Video human.**
  - ➞ **Human feels reality for existence**

Principle of NIRS

- **Near Infra-Red Spectroscopy**
  - **Measurement of the change of quantity of hemoglobin by near infrared rays**
  - **Oxy and deoxy-hemoglobins have different absorption rates**

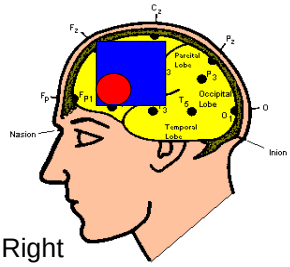


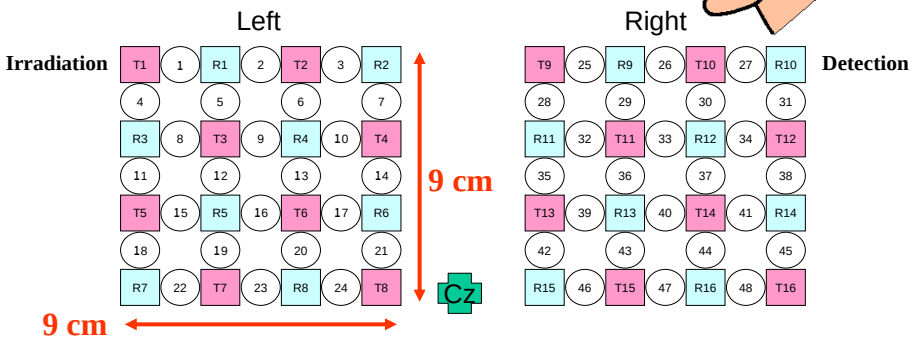


Subjects and probe positions

17 students

(M7, F10, AVE 22.9)



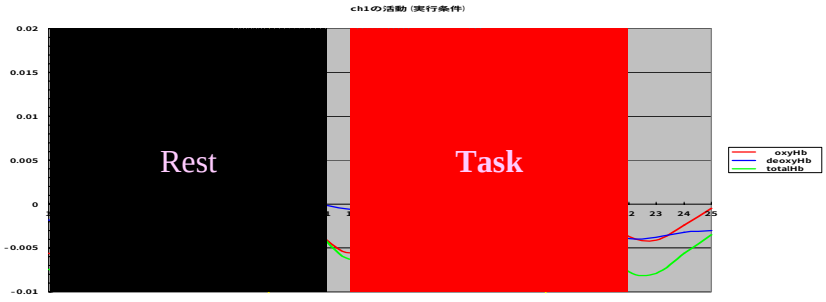


Process of the experimentation

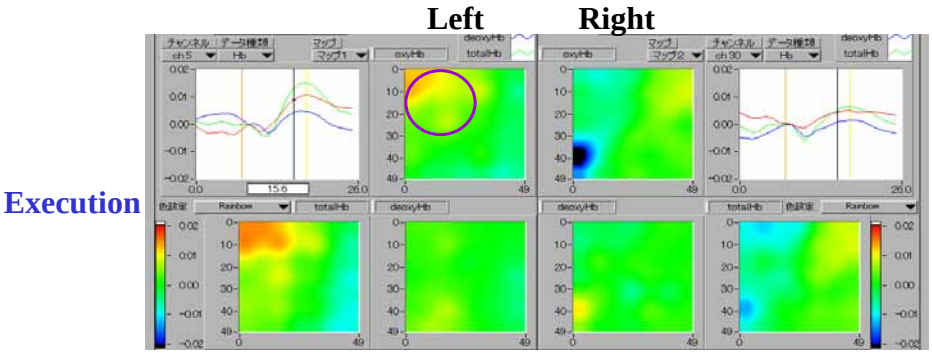
- **Display of Stimuli**
  1. **A human lifting a toy (Real human)**
  2. **A robot lifting a toy (Real robot)**
  3. **A human lifting a toy in video (Video human)**
  4. **A robot lifting a toy in video (Video robot)**
  - **Pre-rest (8 sec.) →Task(10 sec.)→Post-rest (8 sec.)**
  - **Iterate 3 times for each stimulus and display 12 times**
- **The Subject do the task (Execution)**

Data analysis

- **Significance of the signal**
  - **Comparison between “rest” and “task”**
  - **If oxy-hemoglobin increases during “task”, it is recognized as a significant activity**

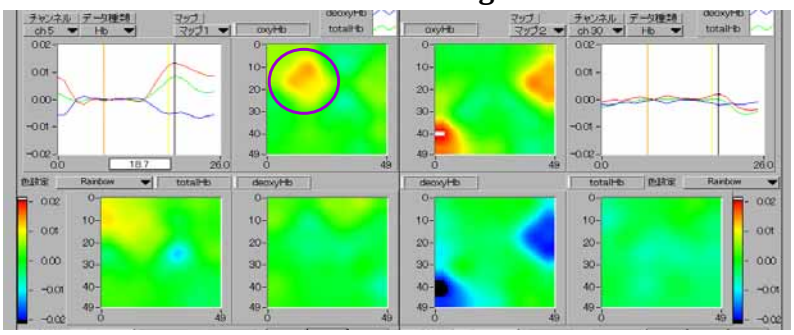


“Real robot” is better than “Video human”?

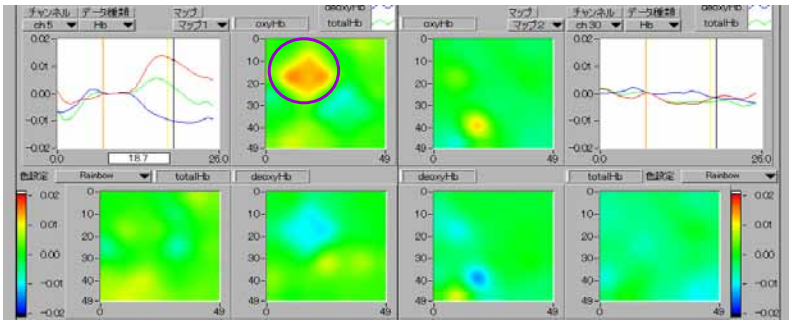


Left Right

Real human

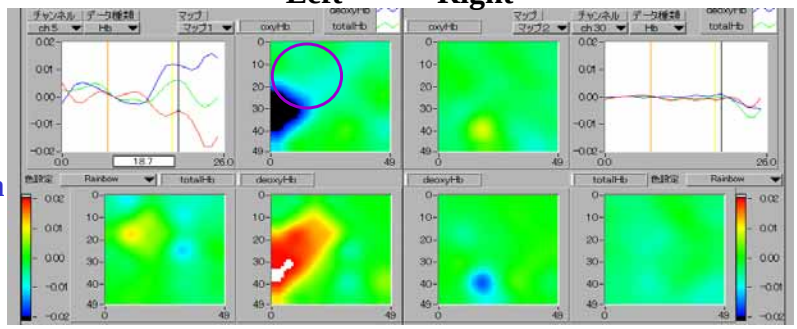


Real robot

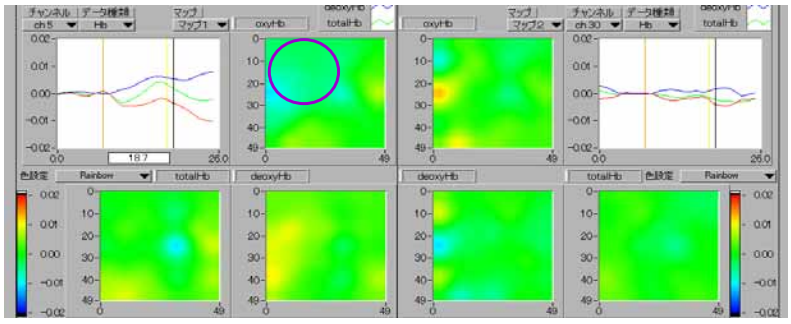


Left Right

Video human



Video robot



Merits and demerits of NIRS

- **Merits (comparing with PET and fMRI)**
  - **Non-invasion**
  - **Real-time measurement**
  - **Compact equipment**
  - **Measurement of two types of hemogrobin**
- **Demerit**
  - **Course spatial resolution**
  - **Limited depth (2-3 cm from the surface)**
  - **Relative measurement**

Fundamental problems

Comparison between robot and human

How does a human distinguish machine, robot and human?
How does the interactive behaviors are developed among them.

Milestones for robot development

Brain activity is an evaluation criteria of the robot performance

References

1. Takayuki Kanda, Takayuki Hirano, Daniel Eaton, Hiroshi Ishiguro, Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial, Journal of Human Computer Interaction (Special issues on human-robot interaction), Vol. 19, No. 1-2, pp. 61-84, 2004.
2. M. Imai, T. Ono and H. Ishiguro, Physical relation and expression: Joint attention for human-robot interaction, IEEE Trans. Industrial Electronics, pp. 636-643, Vol. 50, No. 4, Aug. 2003.
3. A. Arita, K. Hiraki, T. Kanda, and H. Ishiguro, Can we talk to robots? Infants expected interactive robots to be talked by persons, Proc. Biennial Meeting of the Society for Research in Child Development, 2003, Florida, USA.
4. Takayuki Kanda, Takayuki Hirano, Daniel Eaton and Hiroshi Ishiguro, Person identification and interaction of social robots by using wireless tags, Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp.1657-1664, 2003.
5. Takayuki Kanda, Hiroshi Ishiguro, Michita Imai, and Tetsuo Ono, Body movement analysis of human-robot interaction, Proc. International Joint Conference on Artificial Intelligence (IJCAI), pp.177-182, 2003.
6. Itakura, S., Kotani. T., Ishida, H., Kanda, T. & Ishiguro, H., Inferring a robot's false belief by young children. Proceedings of the 32nd Jean Piaget Society, 2002
7. Itakura, S., Ishida, H., Kanda, T., & Ishiguro, H., Robots have false belief, but they do not think false belief task with robot by preschoolers. Hawaii International Conference on Social Science. Proceedings of Hawaii International Conference on Social Science, 2002
8. T. Ono, M. Imai, H. Ishiguro, A model of embodied communications with gestures between humans and robots, Proc. 23th Cognitive Science Society Annual Meetings, pp. 732-737, 2001.