Proceedings of the 2006 IEEE/RSJ
International Conference on Intelligent Robots and Systems
October 9 - 15, 2006, Beijing, China

Comparison of a Humanoid Robot and an On-Screen Agent as Presenters to Audiences

Johane Takeuchi†, Kazutaka Kushida*, Yoshitaka Nishimura*, Hiroshi Dohi*, Mitsuru Ishizuka*
Mikio Nakano†, Hiroshi Tsujino†
*Graduate School of Information Science and Technology, University of Tokyo
7–3–1 Hongo, Bunkyo, Tokyo 113–8656, Japan
†Honda Research Institute Japan Co., Ltd.
8–1 Honcho, Wako, Saitama 351–0188, Japan
Email: johane.takeuchi@jp.honda-ri.com

***Abstract*— Both on-screen agents and humanoid robots are increasingly used as human-computer interfaces. This study evaluates an on-screen agent and a humanoid robot in the task of one-sided presentations. We compared the participant's subjective impressions of nearly identical presentation contents performed by each presenter. The results derived by the Semantic Differential (SD) method and the direct questioning show that each presenter has different functional advantages. We infer that on-screen agent and robot can complement each other in presentations.**

***Index Terms*— humanoid robot, on-screen agent, multimodal presentation, SD method**

I. Introduction

We explore the task of presenting information (using multiple-modalities) to an audience. Artificial agents have already been used in this capacity. For example, consider an animated character that acts and speaks as an attendant to a guest on the web of a shopping site. Recently, humanoid robots have attracted attention to their potentially universal functionality. We are exploring the potential of a humanoid robot for the multi-modal interface. This paper's contribution is the analysis of a humanoid robot and artificial agent as presenters of multimodal information.

While studies of animated characters have been conducted extensively (e.g. [1]), fewer studies have been carried out on humanoid robots as an interface agents [2]–[4]. Kanda *et al.* extensively studied impressions of humanoid robots during human-robot interactions [5]–[7]. They introduced the Semantic Differential (SD) method [8] to evaluate the impressions quantitatively. The studies such as [9],[10] compared a robot with an animated character on screen using psychological evaluation. However, they reported only the effect of physical embodiments of the robot. Additionally, we previously made a preliminary report on the difference of the impression between a humanoid robot and an on-screen agent [11]. However, participants of the experiment were exclusively male professionals from information science.

In this paper, we focus on the different impressions for a presentation given by an on-screen agent and a humanoid robot. We try to determine clear advantages of not only a humanoid robot but also an animated agent on the screen by analyzing the results of psychological experiments. Participants in this experiment are people recruited from the general public. An extensive discussion is provided concerning the complementary characteristics of the two agents.

Clarifying each advantage brings us a design principle for a novel collaborating media interface that uses both a humanoid robot and an on-screen agent. One may imagine a situation where a humanoid robot guides guests in a museum and on-screen agents jointly present details of each item. It is generally said that news shows presented by two or several persons are more comprehensive for audiences than news programs by a single announcer. This result inspires our ongoing work for developing the script language system for joint presentations by an on-screen agent and a humanoid robot.

II. Multimodal Presentation Markup Language (MPML)

In our experiments, MPML (Multimodal Presentation Markup Language) systems are used. In this section, we present an overview of these system. Scripting languages can ease the writing of multi-modal contents with character agents. Several scripting languages have been developed [1]. Among them, MPML [12],[13] was designed for creating multi-modal presentation content. MPML allows many non-specialists to easily write multimodal presentations of life-like character agents. Original MPML targeted only presentations using character agents on a screen monitor, such as the Microsoft Agent [14]. MPML is a medium level description language that does not depend on particular browsers or agent systems.

We have extended MPML to make multi-modal presentation content using humanoid robots [11],[15]. This version of MPML is named as MPML-HR, where HR stands for humanoid robots. We implemented MPML-HR on the humanoid robot Honda ASIMO. As a result, an MPML environment was obtained, in which one can easily write presentation scripts for both ASIMO and the Microsoft Agent, respectively, using almost the same script language.

MPML is an XML (eXtensible Markup Language) based scripting language. MPML has several agent control functions such as position, movements, gestures and the emotion of the

1-4244-0259-X/06/$20.00 ©2006 IEEE

MPML source

```
<mpml>
  <head>
    <title>MPML-HR SAMPLE SCRIPT</title>
      <agent id="PD" character="asimo" />
  </head>
  <body>
    <page ref="C:sample.ppt#1">..... 1
      <move x="1" y="0" />..... 2
      <play act="Wave" />.....3
      <emotion type="happy-for">.....4
      <speak>
        Hello, everyone. My name is ASIMO.  5
      </speak>
     <speak>
          I'm very glad to see you.
      </speak>
      </emotion>
      <point x="50" y="100" />..... 6
      <speak>
        I introduce the function of MPML.
      </speak>
    </page>
  </body>
</mpml>
```

Fig. 1. A sample script of MPML-HR. (1) <page> tag specifies the file of a presentation material on the screen. (2) <move> tag is the command to walk to the specified location in the real space. (3) <play> tag specifies the command of gestures. (4) <emotion> tag specifies the emotion of the agent and modifies the parameters of the speech synthesizer such as tempo, speed, and pitch. (5) <speak> tag specifies the spoken line. (6) <point> tag specifies the pointing action of the robot.

agent. These functions are implemented as XML tags, which is the same as in MPML-HR. Figure 1 shows a sample script in MPML-HR.

III. Evaluation Method

For evaluating each impression of the presentations, we used the MPML-HR for ASIMO and the original MPML for the on-screen agent. Figure 2 shows each agent. Honda ASIMO is a biped humanoid robot. It stands 1200 mm tall and has a mass of 52 kg. ASIMO is similar in height to a human child. It implements several kinds of pre-defined behaviors like dance, greet, guide, nod and so on. One of the Microsoft Agents, "Peedy" [14] was used for the on-screen agent. This is a bird-shaped agent having about sixty gestures. The speech synthesizers for ASIMO and "Peedy" are different. FineVoice, available from NTT-IT, is used in ASIMO. The voice of Peedy is synthesized by Microsoft Speech that is standard for Microsoft Agents.

Each agent presented a weather forecast. We think that the impressions of weather forecasts are less affected by participants' backgrounds while other topics such as politics, sports, culture might be affected more. The scripts of the

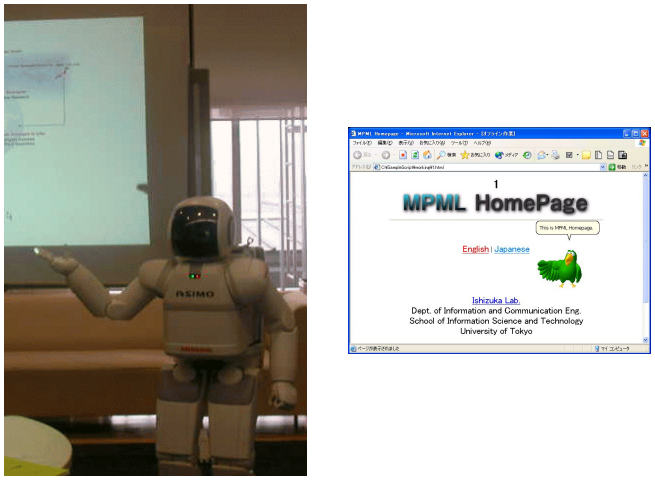


Fig. 2. Two presenters, Honda ASIMO (left), and the on-screen agent (right).

spoken lines and the number of gestures were the same for both ASIMO and the on-screen agent. Each presentation took about 5 minutes. One of the purposes of this experiment was to clarify the advantage of each presentation. We did not make the shape and actions of the on-screen agent equal to those of ASIMO. The most significant difference between the MPML-HR and the original MPML was the pointing action of each agent. In the on-screen agent, the agent can point to a position on the screen by moving to that point while the humanoid robot moves in the real world. If the point operation is executed for ASIMO, the robot walks to one side (right side or left side) of the screen, and then points with his hand. In this experiment, ASIMO walked in front of the screen twice during the presentation, while the on-screen agent moved to point items on the screen many times.

The number of participants in the evaluation was 31. The age of the participants ranged from early-twenties to mid-forties, and the number of males and females was almost the same. The participants were divided into 3 groups for evaluation: group A had nine participants (5 males, 4 females), group B had ten participants (5 males, 5 females), and group C had twelve participants (5 males, 7 females). The participants of group A were shown the presentation of ASIMO first, and the presentation of the on-screen agent second. In groups B and C, the order of presentations was reversed. The on-screen agent's presentation was first, followed by ASIMO's presentation. The participants filled out all questionnaires immediately after each presentation. The results of two participants are omitted due to a lack of descriptions. This left 29 participants results to be analyzed. During the experiment of group B, the system once stalled at the beginning so we mainly used the results of groups A and C for the following analysis. Between the two presentations, the participants were required to carry out a mundane task not associated with the experiment. The total time for completion including breaks was almost two hours.

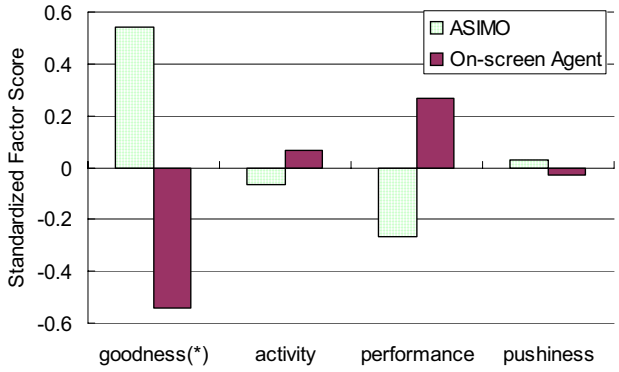
The questionnaires consisted of three parts. The first part was for the SD method. We used 30 pairs of adjectives consisting of a positive and a negative adjective. The par-

ticipants rated each pair of adjectives within seven scales. For example, in the case of a good-bad pair, participants chose among highly good, very good, good, neither good nor bad, bad, very bad, highly bad. The second part consisted of a questionnaire containing ten direct questions. The ten questions were also rated questions in one-to-seven scores. For example, when a participant understood the presentation very well, he/she circled at the highest scores for the question "Was the presentation understandable?" The last part required each participant to write their opinion freely.

IV. Results

First, we describe the results from each part of the questionnaires in turn. After that, the combined analysis of the SD method and direct questions are presented.

A. SD Method



S.D. of the Factor Scores

	goodness	activity	performance	pushiness
ASIMO	0.845	0.891	0.916	0.843
on-screen agent	0.834	1.094	1.010	1.135

Fig. 3. The Standardized factor scores for each factor. These are the results of A and C groups. Standardized factor score means that the factor score standardized to an average 0 and a standard deviation 1. (*) means $p < 0.01$, where p means the statistical significance of the comparison.

In the analysis, these answers are expressed by scores ranging from one to seven. We assigned seven to the most positive expression (e.g. highly good) and one to the most negative one (e.g. highly bad). Factor analysis was performed on the SD method ratings for the 30 adjective pairs. Based on the difference in eigenvalues, we adopted a solution that consists of four factors. Table I shows the retrieved factor matrix which was rotated by a Varimax method. The factor matrix has factor loadings that measure the correlations of each pair of adjectives and factors. The original questionnaires were written in Japanese. The original meanings of Japanese adjectives might be somewhat different from English adjectives shown here.

Factor I seems to represent *goodness* because pairs of adjectives such as 'likable', 'good', 'pretty', and 'favorite' have high absolute values of factor loadings. Since pairs of adjectives such as 'dynamic', 'cheerful', and 'showy' have high absolute values of factor loadings, we name factor II as *activity*. Factor III is named as *performance*, since pairs of adjectives such as 'rapid', 'sensible' bear heavily on performance. Factor IV seems to represent *pushiness*. Table I also shows the evaluated ratings of each adjective pair. For this analysis, we used the results of group A and C for valancing the presentation orders. 'Intelligent' is the most significant different and the higher score is for ASIMO. In the other adjective pairs where there is significant difference, ASIMO is evaluated more positively.

The figure 3 shows the standardized factor score. Only the *goodness* factor is significantly different. The presentation by ASIMO is generally impressive for audiences. T-test on the other factors did not yield significant differences.

B. Direct Questions

TABLE II
Comparison of Presentations

	on-screen agent	ASIMO	Significance
Comprehensible	**4.85**(1.492)	4.35(1.682)	N.D.
Intensity of concentration	**4.95**(1.284)	4.35(2.104)	N.D.
Intensity of interest	4.75(1.479)	**5.70**(1.308)	**5%**
Tempo	3.70(1.187)	**4.35**(1.621)	N.D.
Intensity of impressions	4.55(1.627)	**5.55**(1.161)	**5%**
Suitableness of emotional expressions	**3.70**(1.926)	3.60(1.497)	N.D.
Suitableness of behaviors	**4.25**(1.757)	3.00(1.483)	**5%**
Accurateness of pointing	**4.80**(1.435)	4.05(1.359)	**5%**
Humanlike	3.20(1.778)	**3.85**(1.682)	N.D.
Whole presentation	4.30(1.520)	**4.8**(1.503)	N.D.

The results of the questionnaires of direct questions. These are the results of A and C groups. The mean values are described for the each agent and S.D. values are shown in parenthesis. The significances are the significant differences resulting from a t-test.

Table II shows the results of direct questions. The mean, standard deviation of the scores and statistical significance of the comparison are shown in Table II. The total number of questions where ASIMO's presentation got a higher score, is equal to the on-screen agent's.

ASIMO was evaluated higher in the following questions, "Were you interested in the presentation?," "How do you feel about the tempo of the presentation?," "Did you get a favorable impression from the presentation?," "Was the presentation similar to a human's?," and "Do you think the presentation was good?" Scores are significantly different in the questions about tempo of the presentation and intensity of impressions.

To the contrary, the on-screen agent's was evaluated higher in the following questions, "Was the presentation understandable?," "Did you concentrate on the presentation?," "Were the presenter's emotional expressions appropriate?," "Were the presenter's gestures appropriate?," and "Could the presenter point at the objects accurately?" Scores are significantly different for the questions about the suitability of gestures and pointing accuracy.

C. Free Descriptions

Table III shows the opinions which accounted for the majority. In this analysis, we used all the data from the participants (group A, B, C) to get various opinions. We found

TABLE I

FACTOR ANALYSIS AND SCORES

Adjective Pairs		Varimax Rotated Factor Matrix					Mean and S.D. of Scores		
		Factor I	Factor II	Factor III	Factor IV	Communality	ASIMO	on-screen agent	Significance
Likable	Dislikable	**-0.78675**	0.31936	-0.12703	0.08302	0.74400	**5.56**(1.117)	4.72(1.044)	**5%**
Good	Bad	**-0.78656**	0.51389	-0.08222	0.03956	0.89109	**5.44**(1.117)	4.44(1.212)	**5%**
Pretty	Ugly	**-0.75635**	0.32510	-0.03326	-0.26540	0.74930	**5.94**(0.911)	4.67(1.247)	**1%**
Favorite	Unfavorite	**-0.73486**	0.41393	-0.24401	-0.15967	0.79639	**5.22**(1.083)	4.22(1.133)	**5%**
Interesting	Boring	**-0.73164**	0.35448	-0.15871	0.22901	0.73859	**5.83**(1.302)	4.44(1.461)	**5%**
Superior	Inferior	**-0.72644**	0.40691	-0.20623	0.07246	0.74107	**5.33**(1.000)	3.83(1.067)	**1%**
Intelligent	Unintelligent	**-0.71394**	0.09645	0.02644	0.23454	0.57472	**5.22**(1.030)	3.94(1.079)	**0.1%**
Full	Empty	**-0.68114**	0.37493	-0.10571	-0.00182	0.61570	**4.89**(1.197)	3.61(1.061)	**1%**
Considerate	Selfish	**-0.61707**	0.07231	-0.31953	-0.40729	0.65399	**4.44**(1.012)	3.78(1.030)	N.D.
Pleasant	Unpleasant	**-0.58344**	0.46376	-0.20457	-0.10475	0.60829	**5.39**(1.208)	4.67(1.291)	N.D.
Intelligible	Unintelligible	**-0.57532**	-0.00422	-0.05391	-0.02164	0.33438	4.06(1.615)	4.06(1.471)	N.D.
Warm	Cold	**-0.55837**	0.44342	-0.19611	-0.24736	0.60805	**4.56**(1.212)	4.11(1.197)	N.D.
Exciting	Dull	**-0.53243**	0.51019	0.08630	0.02465	0.55183	**5.50**(1.258)	4.38(1.208)	**5%**
Calm	Agitated	**-0.52698**	0.20534	-0.09974	-0.24078	0.38780	4.89(1.410)	**5.00**(1.106)	N.D.
Humanlike	Mechanical	**-0.52209**	0.42453	-0.23427	-0.10680	0.51909	**3.94**(1.580)	3.33(1.291)	N.D.
Friendly	Unfriendly	**-0.51876**	0.39978	-0.17959	-0.07094	0.46622	**4.56**(2.034)	4.44(1.343)	N.D.
Distinct	Vague	**-0.51790**	0.47188	-0.02801	0.07356	0.49709	4.67(1.155)	**4.72**(1.283)	N.D.
Dynamic	Static	-0.21460	**0.77320**	0.13463	0.14000	0.68161	**5.06**(1.129)	4.78(1.133)	N.D.
Cheerful	Lonely	-0.30707	**0.75108**	-0.24454	-0.12478	0.73379	**4.78**(0.975)	4.56(1.343)	N.D.
Showy	Quiet	-0.11537	**0.71180**	-0.06622	0.33309	0.63531	**4.16**(1.014)	3.94(1.268)	N.D.
Light	Dark	-0.28257	**0.70948**	-0.22059	-0.24463	0.69171	4.56(1.212)	**4.89**(1.370)	N.D.
Frank	Rigid	-0.41596	**0.60230**	-0.09894	0.03290	0.54666	4.72(1.660)	**5.11**(1.149)	N.D.
Lively	Lifeless	-0.52460	**0.55996**	-0.15679	-0.02109	0.61379	**4.67**(1.333)	3.61(1.533)	N.D.
Active	Passive	-0.27835	**0.50695**	-0.21912	0.21875	0.43034	**4.61**(1.061)	4.44(1.165)	N.D.
Kind	Cruel	-0.26043	**0.44639**	-0.33811	-0.32290	0.48567	**4.77**(1.397)	4.72(1.145)	N.D.
Rapid	Slow	-0.03381	0.09386	**-0.75667**	0.10904	0.59440	**3.63**(1.086)	3.61(1.112)	N.D.
Quick	Slow	-0.18258	0.00038	**-0.62056**	0.34929	0.54044	3.28(1.407)	**3.83**(1.213)	N.D.
Sensible	Insensible	-0.22581	0.40607	**-0.51234**	-0.04253	0.48018	3.50(0.764)	3.50(1.067)	N.D.
Aggressive	Timid	0.04296	0.12938	-0.11519	**0.66086**	0.46858	4.22(0.916)	**4.61**(1.112)	N.D.
Complicated	Simple	-0.01391	-0.00324	-0.09987	**0.54571**	0.30798	**3.67**(1.202)	3.44(1.165)	N.D.

The results of the questionnaires for the SD method. The factor matrix received by factor analysis of the results by all participants and Varimax rotation. The means and standard deviation (S.D.) of the ratings of 18 subjects (A and C group), for 30 adjective pairs. The mean values are provided for the each agent and S.D. values are shown in parenthesis. The significances column lists the significant differences resulting from a t-test.

TABLE III

MAJORITY OF OPINIONS

	proportions
Couldn't concentrate on presentation due to curiosity to ASIMO	61.3%
About voice quality	83.9%
About gestures	80.6%
About emotional expressions	32.3%

The largest proportion of opinions from the free description questionnaire. These are the results of A, B and C groups.

many complaints about the voice qualities of the synthesized speech. We used many gestures in the both presentations, intending to make the presentation more attractive. Most of the participants thought that less gestures were preferable in the presentation of ASIMO.

Opinions about emotional expressions of robots were divided into two types. Some participants thought that the emotional expressions of ASIMO through gestures should be more exaggerated because ASIMO does not have facial expressions. On the contrary, some participants thought that emotional expressions were not necessary in this case. It has been hypothesized that somewhat exaggerated emotional expressions in content using on-screen agents improve naturalness. We could not confirm whether exaggerated emotional expressions were effective in the case of ASIMO.

Some participants thought that the pointing actions of ASIMO were inferior to those of the on-screen agent. We found that many participants wrote opinions such as "I could not concentrate on the presentation itself due to my heavy interest in ASIMO."

D. Combined Analysis of Impressions and Evaluations

Some participants evaluated the presentation of ASIMO better than that of the on-screen agent and some were vise versa. By summing the scores of the ten direct questions, we separated the participants into a group that preferred ASIMO and a group that preferred the on-screen agent. The results are depicted in Fig. 4. The agent group has more widely distributed scores than those of the ASIMO group. The participants were equally divided between the ASIMO and Agent group.

We investigated the tendencies of each group in the SD method. Figure 5 shows the results in factor analysis. The participants of the ASIMO group tend to evaluate ASIMO predominantly in the *goodness* factor, which represents general

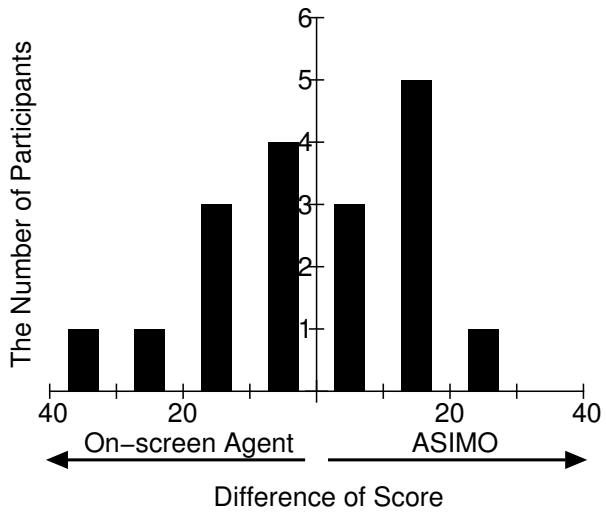


Fig. 4. The distribution of the difference of scores (which sum all the scores of direct questions). This result is from the 18 participants of A and C groups.

impressions. On the other hand, the Agent group tends to evaluate the on-screen agent in the factors of *performance* and *pushiness*, which are represent performance of the presentation. The score differences of Agent group is generally less significant than that of the *goodness* factor in the ASIMO group. Agent's group evaluated the presentations by both presenters as almost the same with regard to impression.

V. Discussion

The real presence of the biped humanoid is impressive to audiences. The most definitive result of comparing ASIMO and the on-screen agent is that abstract adjectives such as intelligence, likableness have significant differences. This was also indicated in our previous report [11], in which the evaluation was participated by professionals. These kinds of impressions are also important in presentations of everyday life. Even for the participants who evaluated the on-screen agent better than ASIMO, the general impression of ASIMO can be favorable or almost equal to the on-screen agent. These impressions are thought to be important not only in presentations but also in interactions with humans.

As a result,

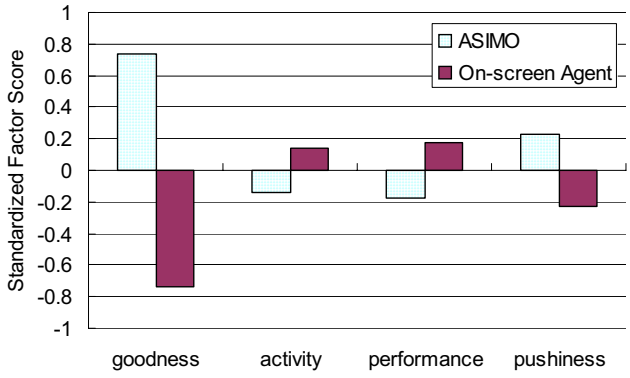
- A humanoid robot ASIMO can be a presenter instead of on-screen agents.

On the other hand, our results also suggests that

- it is necessary to make an audience pay more attention to the presentation materials on the screen and
- to improve the pointing ability and the expressions of gestures.

Several participants evaluated the on-screen agent better than ASIMO. They basically emphasized on the performance of the presentations rather than their impressions. On-screen agents such as the Microsoft Agent are generally well designed for presentations like those described in this paper and for user interfaces of computers. This suggests that the humanoid

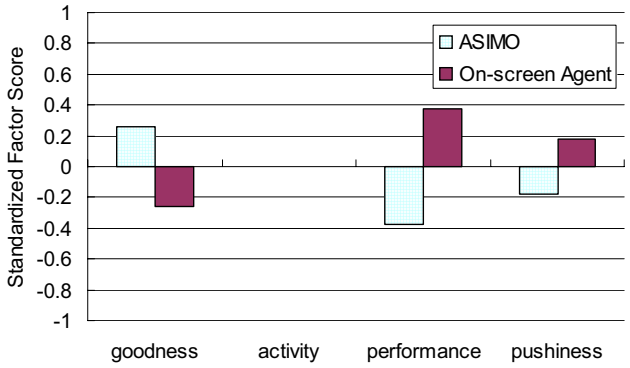
(a) ASIMO group



S.D. of the Factor Scores

	goodness	activity	performance	pushiness
ASIMO	0.729	0.593	0.972	0.799
on-screen agent	0.613	1.269	0.996	1.121

(b) On-screen agent group



S.D. of the Factor Scores

	goodness	activity	performance	pushiness
ASIMO	1.027	1.080	0.804	0.877
on-screen agent	0.900	0.914	1.037	1.079

Fig. 5. The factor analysis of groups pooled according to preference for ASIMO or on-screen agent.

robots and on-screen agents can be complementary. Regardless of future developments of humanoids, on-screen agents will be maintaining their advantages in pointing and speed of motions. On-screen agents have the advantage of attracting audience attention to the screen. When an on-screen agent moves to the object on the screen and explains it, the audience follows the agent with their eyes and watches the object being explained.

A. Generality of the Results

We used Microsoft Agents for a on-screen agent, which we can make common scripts to ASIMO by means of MPML. There are already more attractive on-screen agents than the Microsoft Agents. The score of 'goodness' factor can be comparable to that of ASIMO if we use such an attractive on-screen agent. Still we can assert that ASIMO is impressive

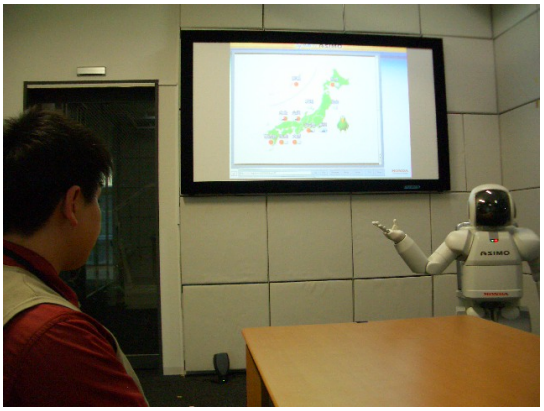


Fig. 6. The collaborative presentation by ASIMO with the animated agent on the screen.

to audiences in absolute values. This paper does not deny on-screen agents. We try to deduce the characteristics and properties of humanoid robots represented by ASIMO and on-screen agents represented by the Microsoft Agent. ASIMO has the advantage in impressions to audiences even in tasks where on-screen agents can achieve better performances.

B. For Making Good Presentation Contents

Humanoid robots and on-screen agents have both advantages and disadvantages as a presenter. The advantages of a humanoid robot are:

- because it has a real body and its size and shape are close to a human's, it can give an intense impression in a presentation,
- it can give not only a presentation using slides but also a presentation about real objects, and
- since it can move in real space, approach and gaze at the audience, it can give a sense of unity with the audience in a presentation.

The disadvantages are:

- because of the limitation in raising hands, it is difficult to point at objects accurately,
- audiences find it somewhat difficult to pay attention to the screen.

On the other hand, an on-screen agent has the following advantages:

- it can accurately point at objects it is explaining because it exists in the display and moves there,
- it can attract the audience's attraction into slides and explained objects.

Their disadvantages are:

- it has a difficulty in explaining real objects outside of the display, and
- it is less impressive to audiences.

Considering their advantages and disadvantage, we think that a humanoid should be a main presenter and an on-screen agent should compensate for its lack of abilities to point at objects accurately and attract an audience's attention to the screen. We think that when we design a presentation in this way, the presentation gives intense and favorable impressions to audiences. We think that audiences not only have a sense of unity with the presenters but also pay attention to slides.

We are extending MPML-HR to MPML-HR2. In the MPML-HR2 system, we can use both a humanoid robot and an on-screen agent effectively. They perform a presentation supporting each other (Fig. 6).

ACKNOWLEDGMENT

The authors would like to thank Yuji Hasegawa, Tokitomo Ariyoshi and Kazuhiro Nakadai for helping the psychological experiment and Carson Reynolds for comments on the manuscript.

REFERENCES

[1] H. Prendinger and M. Ishizuka, Eds., *Life-Like Characters: Tools, Affective Functions and Applications*. Springer-Verlag, 2004.

[2] A. Hoshino, K. Kato, J. Takeuchi, and H. Tsujino, "A chat information service system using a humanoid robot," in *Proceedings of 14th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN'05)*, 2005, pp. 468–473.

[3] H. Asoh, N. Vlassis, Y. Motomura, F. Asano, I. Hara, S. Hyamizu, K. Ito, T. Kurita, T. Matsui, R. Bunschoten, and B. Kröse, "Jijo-2: An office robot that communicates and learns," *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 46–55, 2001.

[4] M. Nakano, Y. Hasegawa, K. Nakadai, T. Nakamura, J. Takeuchi, T. Torii, H. Tsujino, N. Kanda, and H. G. Okuno, "A two-layer model for behavior and dialogue planning in conversational service robots," in *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05)*, 2005, pp. 1542–1548.

[5] T. Kanda, H. Ishiguro, and T. Ishida, "Psychological analysis on human-robot interaction," in *Proceedings of 2001 IEEE International Conference on Robotics and Automation (ICRA'01)*, 2001, pp. 4166–4173.

[6] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu, "Development and evaluation of an interactive humanoid robot "Robovie"," in *Proceedings of 2002 IEEE International Conference on Robotics and Automation (ICRA'02)*, 2002, pp. 1848–1855.

[7] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, and H. Ishiguro, "Analysis of humanoid appearances in human–robot interaction," in *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05)*, 2005, pp. 62–69.

[8] J. G. Snider and C. E. Osgood, *Semantic differential techniques, a sourcebook*. Chicago: Aldine, 1969.

[9] C. D. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in *Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, Sep. 2004, pp. 3559–3564.

[10] K. Shinozawa, F. Naya, J. Yamamoto, and K. Kogure, "Differences in effect of robot and screen agent recommendations of human decision-making," *International Journal on Human–Computer Studies*, vol. 62, no. 2, pp. 267–279, 2005.

[11] Y. Nishimura, k. Kushida, H. Dohi, M. Ishizuka, J. Takeuchi, and H. Tsujino, "Development and psychological evaluation of multimodal presentation markup language for humanoid robots," in *Proceedings of 5th IEEE-RAS International Conference on Humanoid Robots (Humanoids'05)*, 2005, pp. 393–398.

[12] "Mpml home page." [Online]. Available: http://www.miv.t-utokyo.ac.jp/MPML/mpml.html

[13] H. Prendinger, S. Descamps, and M. Ishizuka, "Mpml: a markup language for controlling the behavior of life-like characters," *Journal of Visual Languages and Computing*, vol. 15, no. 2, pp. 183–203, 2004.

[14] "Microsoft agent home page." [Online]. Available: http://www.microsoft.com/msagent

[15] K. Kushida, Y. Nishimura, H. Dohi, M. Ishizuka, J. Takeuchi, and H. Tsujino, "Humanoid robot presentation through multimodal presentation markup language mpml–hr," in *Proceedings of 4th International Joint Conference on Autonomous Agents and Multi Agent System (AAMAS'05) Workshop 13 "Creating Bonds with Humanoids"*, 2005, pp. 23–29.