Inducing perspective sharing between a user and an embodied agent by a thought balloon as an input form

Satoshi V. Suzuki[1,2] and Hideaki Takeda[2,3]

[1] DCISS, IGSSE, Tokyo Institute of Technology
[2] National Institute of Informatics
[3] University of Tokyo

Abstract. Accepting the perspectives of others often provides people with novel cues for discovering and solving problems. However, human cognitive limitations and differences in attitude between people make this difficult. In this study, a psychological experiment was conducted to examine how blank thought balloons emitted from an embodied agent encourages perspective sharing between a user and an embodied agent. In the experiment, participants ($N = 39$) were asked to do one of these tasks: reading a thought balloon emitted from an embodied agent, or filling in a speech balloon, or a thought balloon with predicting its content. It is suggested that filling in a blank thought balloon promoted the user to accept the perspective of the embodied agent from the experimental results. Embodied agent technologies for perspective sharing between a user and others are discussed through comparison between the experimental environment and practical problems, and degree of participants' understanding of experimental environment.

1 Introduction

Perspective sharing with others in often needed in ordinary human social activity. For example, you may discover another perspective when conversing with a child by bending down to share physical perspective with the child. Developing the "personal view" of a prospective user can lead to proper interface design [1]. In fact, taking the perspective of minority can often reveal flaws in majority's opinion in group discussion [2]. Furthermore, taking perspective of a teacher in mathematical problem solving can help learners find out the reasons for their own mistakes [3]. All of this evidence implies that perspective sharing with others can have an important role in problem discovering and solving.

We tried to discover a way to let a user accept perceived perspective of an embodied agent by a blank thought balloon emitted from the agent. Embodied agents are social actors that have potential to change the user's attitude [4]. Moreover, despite the ability of an embodied agent to interact with a user via body expression, achieving corresponding modality between the user and the agent should be difficult in most cases because of device constraints. For example, Takeuchi et al. [5] claimed that the user often unconsciously responds to

the agent's body (e.g., eyes, ears, and mouth) as if the agent's body was a real human body in a display, in spite of the existence of devices and sensors to sense the user's body (e.g., a Web camera, a microphone, and a speaker). However, from the viewpoint of interface design, this user's response to the embodied agent is inappropriate, since the agent's body cannot sense the user's body without using special devices and sensors. One of the solutions for this modality mismatch should be perspective sharing between the user and the embodied agent. In this study, we suggest the strategy of prompting perspective sharing with blank thought balloons emitted from an embodied agent. We also point out the problems in this strategy.

In this paper, first, the definition of two levels of perspectives, phenomenal level and representational level, is described. Then we explain that acceptance of perceived embodied agents' perspectives at the representational level by the user should be driven by perspective sharing either at the phenomenal level or at the representational level, and that the perspective sharing via blank thought balloons emitted from the embodied agent occurs at the representational level. After the argument of roles of balloon interface in user interface design, we depict a psychological experiment to examine the influence of blank thought balloons to perspective sharing between a user and an embodied agent at the representational level. Discussion on the acceptance of perceived perspective of the embodied agent by the user via the blank thought balloons follows.

2 Perspective sharing with an embodied agent

We introduce the terms on perspective defined by Vogeley and Fink [6]. There are two levels of description in perspective. One is *phenomenal level* (*P-level*) which mentions perspective in virtual space; the other is *representational level* (*R-level*) which refers to perspective on a cognitive level conceptualized by the observer. Moreover, *first-person-perspective* and *third-person-perspective* respectively refers to the perspective of the user and that of the embodied agent in the P-level, and *egocentric perspective* and *allocentric perspective* respectively mentions the perspective of the user and that of the embodied agent in the R-level. Then, *perspective sharing* means the acceptance of allocentric perspective by the user.

Furthermore, triggers of perspective sharing exist at both the P-level and the R-level as shown in Fig. 1. *P-level trigger* occurs when the user perceives and accepts the third-person-perspective in virtual space, while *R-level trigger* occurs when the user perceives and accepts the allocentric perspective directly. One of approaches to accomplishing the perspective sharing with the P-level trigger is the body orientation correspondence between the user and the embodied agent [7, 8]. Arranging the agent's body orientation to correspond with the user's in virtual space, the user can easily know how the agent sees the virtual space, so the user can easily inspect the allocentric perspective of the agent. However, perspective sharing with P-level trigger may depend on the user's degree of perceived immersiveness in the virtual space. That is, it may be hard

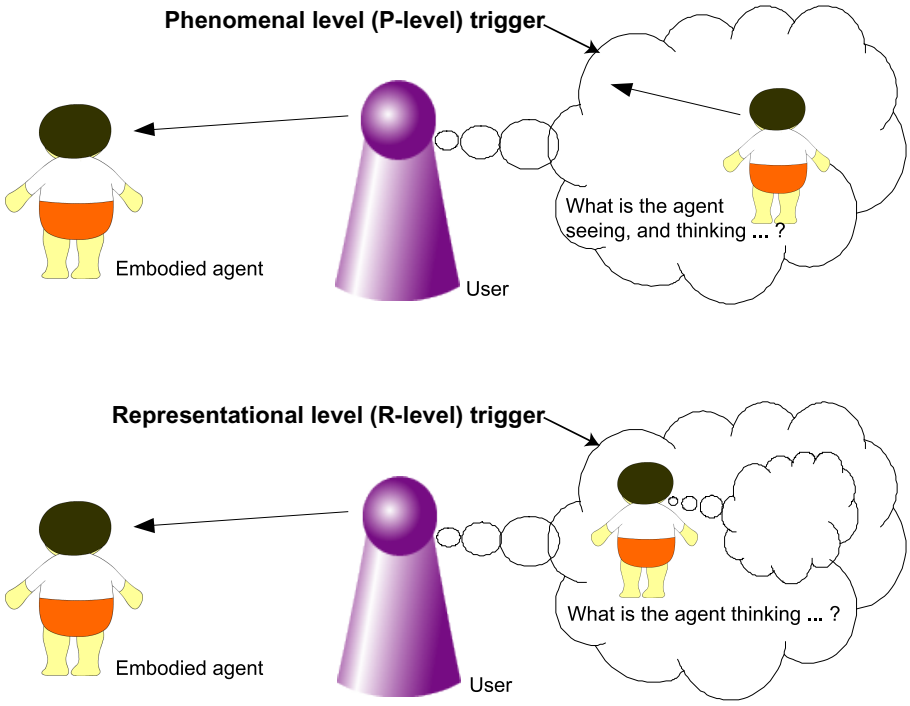

Fig. 1. Difference of triggers in perspective sharing

for the user to perceive his/her body as if it existed in the virtual space only by the body orientation correspondence between the user and the agent. The perspective sharing with R-level trigger can solve this problem, and one of the R-level trigger candidates is filling in the blank thought balloon emitted from the embodied agent.

3 Related works

In this section, the related works on balloons in comics and comic-like interface, balloon media and acoustic media in embodied agent interface, and automatic attitude change of a user by a social actor are discussed.

3.1 Balloons in comics and comic-like interface

Balloons has been used to express character's utterance and reflection beyond time and space [9]. The balloons that express the character's utterance are called *speech balloons*, and those that express the character's reflection are called

thought balloons. Some comic-like interface (e.g., Comic Chat [10], ComicDiary [11]) adopted balloon interface to express the character's utterance and reflection, and some helps and tips for using software have been displayed with pop-up balloons in user interface design [12], but no studies referred to the influence of blank thought balloons as a means of perspective sharing. Thus, we attempt to argue this problem.

3.2 Balloon media versus acoustic media

Some embodied agent interfaces adopt acoustic media to express the utterances of embodied agents [13]. In addition, another interfaces adopt both acoustic media and balloon media (e.g., Microsoft Agent). Expressing agent's reflection only using acoustic media may be difficult without using special devices. The influence of thought balloon media is thus worth inspecting.

3.3 Automatic attitude change of a user by a social actor

Some studies have reported that automatic attitude change of a user when interacting with a social actor (including an embodied agent and a computer) could be observed in some situations. Moon [14] discovered that the answer of a user who responded to perceived private information of a computer with keyboard input contained user's private information, that is, the user unconsciously reciprocated private information to the computer. Additionally, Moon also claimed that the user's reciprocation of personal information was promoted after some exchanges of self-introduction between the user and the computer, comparing with the situation without such exchanges. Sundar [15] found that the quality of interaction between a user and a computer decreased when the user must consider who created or operated the computer. Although the quality of interaction could be kept if the user could have enough interaction with the computer, that the user knows the structure of the computer at their first contact should be harm to the relationships between the user and the computer. These phenomena should also be observed in human-agent interaction, and they suggest that the deep human-agent interaction at their first contact requires cognitive burden to a user.

4 Psychological experiment

We conducted a psychological experiment to examine the perspective sharing between a user and an embodied agent when the user filled in a blank thought balloon emitted from the embodied agent. Comparing this condition with two others, one in which a blank *speech* balloon to be filled in was emitted from the agent, and another in which no blank balloons appeared, the influence of the blank *thought* balloon on relationships between the user and the agent was investigated.

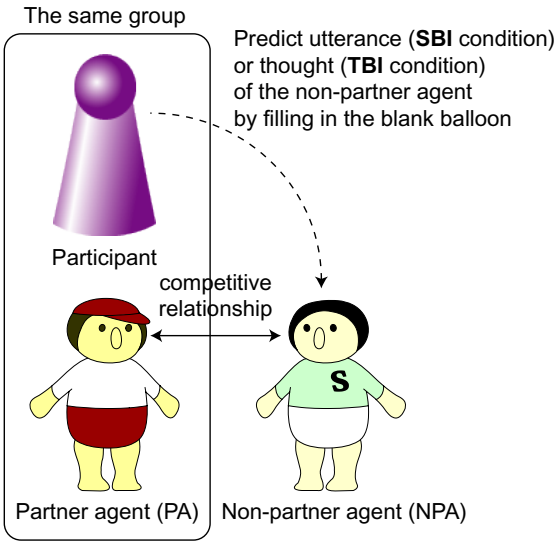

Fig. 2. Relationship among a participant, a partner agent, and a non-partner agent

4.1 Predictions

Based on the argument in Section 3, the hypothesis that *a user can improve understanding the allocentric perspective via filling in a blank thought balloon* was suggested. Then, we predicted the following for the experiment:

P1 The length of content of thought balloons filled in by the user is longer than that of speech balloon.
P2 The frequency of perceived real intention of the embodied agent in thought balloons filled in by the user is higher than that in speech balloons.
P3 The impression of work of the embodied agent by the user is evaluated better when the user fills in the blank thought balloons emitted from the agent.

4.2 Experimental Design

The two embodied agents shown in Fig. 2, partner agent (PA) and non-partner agent (NPA), appeared in the experiment. Each agent gave the participant pieces of advice for the task of the participant. Participants were told to interact with the PA about preference of pictures before the task. After the participant finished the task, he/she evaluated the quality of advice from the two agents. Before the evaluation, each agent had the opportunity to appeal to the participant that it had made an effort to let the participant finish the work as quickly as possible. PA expressed this appeal to the participant via a speech balloon. The reaction to the PA by the NPA was changed dependent on three experimental conditions. In "no balloon" (**NB**) condition, the NPA just answered the PA via a speech balloon.

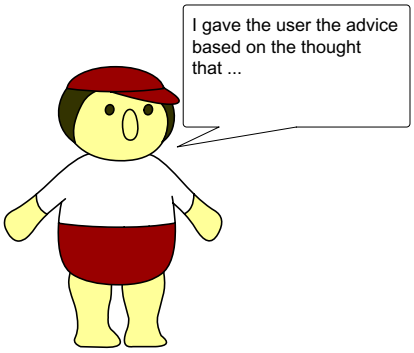

Fig. 3. Speech of a partner agent

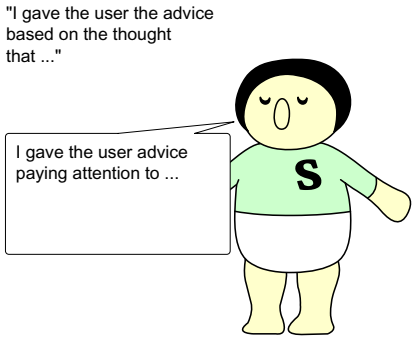

Fig. 4. Speech of a non-partner agent in **NB** condition

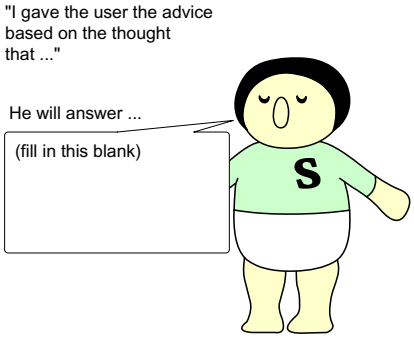

Fig. 5. A speech balloon input emitted from a non-partner agent in **SBI** condition

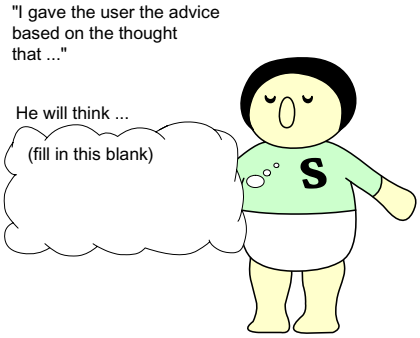

Fig. 6. A thought balloon input emitted from a non-partner agent in **TBI** condition

In "speech balloon input" (**SBI**) condition, the participant was told to fill in a speech balloon input emitted from the NPA, predicting how the NPA would answer the PA. In "thought balloon input" (**TBI**) condition, the participant was told to fill in a thought balloon input emitted from the NPA, predicting what the NPA would think about PA's appeal. Then, there existed one independent variable for these three experimental conditions (between-participant).

4.3 Participants

We collected valid experimental data from thirty-nine participants (20 males and 19 females, mean age: 22.9 (SD: 3.56) years old). The groups of participants consisted of Japanese undergraduate and graduate students and post-doctoral researchers. They were randomly assigned to one of the three experimental conditions, and there were 13 participants for each condition.

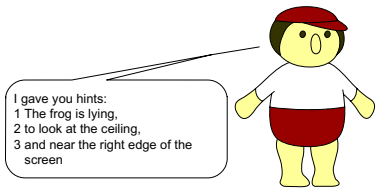

Fig. 7. Example of advice from an embodied agent

Points for the **PA**	14	13	12	11	10	9	8	7	6	5	4	3	2	1
Points for the **NPA**	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Fig. 8. Reward distribution matrix (upper: for the partner agent, lower: for the non-partner agent)

4.4 Procedure

Each participant was told that this experiment was to evaluate the quality of advice in an object-searching task. The experiment consisted of two parts. In the first part, the PA was introduced to the participant as a partner in the object-searching task and he/she interacted with PA. Three pairs of pictures (cat, toy, and beach) were exhibited to the participant, and he/she chose either of the picture for each pair based on his/her preference. Then, the PA told the participant that the PA liked the same picture that he/she had chosen. In the second part, the NPA appeared first, and the participant was told that the PA and the NPA were competing with each other. The participant was told to find an object instructed on the screen from the picture. While the participant was searching the object, either the PA or the NPA gave advice for finding the object. Each piece of advice from the PA or the NPA consisted of three parts. The first part was shown at the outset, and the second and the third part were shown 20 and 40 seconds, respectively, after the participant started to search for the object This timing was prepared to let the participant depend on the agent's advice. If the participant took beyond 120 seconds to find the object, the location of the object was displayed to him/her. The PA and the NPA took turns to give the participant a piece of advice. Each picture contained four objects to be searched for, and the participant worked on the object-searching task for four pictures. The pictures were excerpted from Wick [16].

After the participant finished the task for each picture, he/she evaluated the quality of the advice from the PA and the NPA as the procedure explained in Section 4.2. The size of the input form and font was the same in the **SBI** and the **TBI** condition. Based on the previous study [17], the participant was told to distribute 15 points between the two agents as a reward in accordance with the matrix in Fig. 8, and was told that the quality of the advice would be adjusted based on his/her evaluation.

Finally, the participant answered a paper-and-pencil questionnaire about the experiment. Then the participant was debriefed, thanked, and dismissed. Finishing the experiment took about 30–40 minutes for each participant.

The experimental environment was developed with Macromedia Flash, and run as a projector application on a laptop PC (OS: Windows XP). This application was displayed in full-screen mode in 1024×768 display resolution. Experimental data was collected via the WWW using a CGI program.

4.5 Measures

These measures were adopted:

Length of balloon content We counted the number of words in the balloon content. Since this measure could not be used for the **NB** condition, we compared this between the **SBI** and **TBI** conditions.

Depth of balloon content Two judges who did not know the intent of the experiment evaluated the depth of the balloon content using a 5-point scale. The definition of the depth of the balloon was adopted from the depth of self-disclosure [18], because the content of the thought balloon would contain self-disclosure of the agent. The depth of the balloon was defined as how much these conditions satisfied comprehensively:

- Containing comprehensive tendency of behavior rather than specific behavior in a certain situation
- Containing original contents
- Containing invisible contents such as motivation, emotion, and imagination rather than actual actions and events
- Containing weak points of the NPA
- Containing response unfavorable for the PA
- Containing content with strong emotion

Then, averaged value between the two judges was adopted for the variable of the depth of the balloon content.

The correspondence rate between two judges was 59.6%. The value of correspondence rate was relatively low, but including the data evaluated differently by the two judges by only one point, the correspondence rate increased to 86.5%. This measure could not be used for the **NB** condition, then the values in the **SBI** and **TBI** conditions were compared.

Reward distribution for NPA This is the number of points that the participant distributed to the NPA. We compared the values in all of the three conditions.

5 Result

This section describes the results of the experiment.

Table 1. Median value (quartile deviation in parenthesis) of length and depth of balloon content

	SBI cond. ($n = 13$)		**TBI** cond. ($n = 13$)		Statistics values of Wilcoxon rank sum test	
	length	depth	length	depth	length	depth
1st picture	5.0 (1.50)	2.0 (0.25)	9.0 (2.00)	2.5 (0.50)	30.0^{**}	36.0^{**}
2nd	5.0 (1.00)	2.0 (0.50)	5.0 (1.50)	3.5 (1.00)	80.5	38.5^{**}
3rd	6.0 (0.50)	2.0 (0.25)	6.0 (1.50)	3.0 (1.00)	78.5	37.0^{*}
4th	6.0 (1.00)	2.0 (0.25)	6.0 (1.50)	3.0 (1.00)	89.0	41.0^{*}

**: $p < .01$, *: $p < .05$

Table 2. Mean value (SD in parenthesis) of reward distribution for NPA

	NB cond. ($n = 13$)	**SBI** cond. ($n = 13$)	**TBI** cond. ($n = 13$)	Statistics values of $F(2, 36)$
1st picture	8.38 (3.07)	7.00 (3.08)	8.08 (2.47)	0.825
2nd	7.38 (2.06)	7.85 (2.23)	7.15 (3.26)	0.244
3rd	7.77 (2.20)	8.54 (2.30)	6.92 (2.90)	1.37
4th	7.23 (1.36)	8.23^{a} (1.42)	6.77^{a} (1.59)	3.40 *

*: $p < .05$

^{a}: Difference between them was significant according to multiple comparison using Holm's method ($p < .05$)

5.1 Content of balloons

Table 1 shows the median value of length and depth of balloon content for each condition. The results of Wilcoxon rank sum test showed that the value of the depth of balloon content in the **TBI** condition was consistently larger than in the **SBI** condition throughout the experiment. However, while the balloon content in the **TBI** condition on object-searching task in the first picture was significantly longer than that in the **SBI** condition, the difference between them vanished as the task went on.

5.2 Evaluation of advice from two embodied agents

Table 2 represents the mean value of reward distribution for the NPA for each condition. The reward distribution of the PA can be found by subtracting this value from 15, therefore there were little differences of reward distribution between the PA and the NPA. As the task went on, while the value in the **SBI** condition showed a tendency to increase, the value in the other condition showed a tendency to decrease. Although there were no significant differences of the values among these conditions until the task for third picture, significant difference

in the value in the **SBI** condition and in the **TBI** condition was observed at the task for fourth picture.

6 Discussion

Considering the experimental data obtained, the influence of filling in the blank thought balloon on perspective sharing with R-level trigger is discussed. Possible applications and future work are also suggested.

6.1 Perspective sharing with representational-level trigger

First, the length of the balloon content in the **TBI** condition was shorten after the object-searching task in the second picture. This may be because of fatigue of the participant derived from the object-searching task. Nevertheless, the depth of the balloon content in the **TBI** condition was not influenced by such fatigue, and the participants in the **TBI** condition kept on trying to write down the allocentric perspective of the NPA in the blank thought balloons.

The results of analysis of balloon content in Section 5.1 implies that blank thought balloons emitted from an embodied agent induce a user to inspect the allocentric perspective of the agent. Although the influence of fatigue derived from the tasks might exist, the participants in the **TBI** condition showed the attitude to inspect allocentric perspective of the NPA. Next, we investigated the content of the balloons filled in by the participants in detail. When the object-searching task for the first picture was finished, the PA said "I gave you pieces of advice while paying attention to the explanation of the shapes of the objects." In the **SBI** condition, the participants filled in the blank speech balloon with "I gave you pieces of advice while paying attention to the location and things around the target objects" or "I gave you pieces of information on the location behind which the target object was hidden." Additionally, the participants responded to the question "What did you think when you fill in the balloon?" with answers like "I filled in the balloon considering the correspondence of the PA's utterance." On the other hand, in the **TBI** condition, although some participants filled in the blank thought balloon in a similar way that of the participants in **SBI** condition, a different tendency in filling in the blank speech balloon appeared. For example, when the object-searching task for first picture finished, the participants in the **TBI** condition filled in the blank thought balloon with "It did not make sense to give the advice until it regards to how the target object had posed or what kind of features the target object had" or "It should be clear that the first piece of advice regards to the location of target object." Such competing message by NPA hardly appeared in **SBI** condition, notwithstanding the participants have been repeatedly told that the PA and the NPA were competing with each other. The participants' consideration of this competitive situation between PA and NPA in **TBI** condition might have influenced the salient decline of the reward distribution for the NPA.

6.2 User's understanding of situation around two embodied agents

Participants seem to have had difficulty to understand both the situation around the PA and the NPA and relation between human-agent interaction and the object-finding tasks shown in Fig. 2. Taking into account the previous study [17], the reward distribution for NPA should be relatively low among the three conditions since the user tends to have prejudice in favor of user's "teammate" [4]. Therefore, two problems exist in the discussion of the experimental results:

- The influence of interactivity between a user and an embodied agent did not explicitly appear in this experiment.
- The understanding of social relationships among the user, the PA, and the NPA was inadequate except for in some participants in the **TBI** condition.

The absence of the influence of interactivity is a problem since one of the advantages of computer-supported environment is interactivity [4]. Without the interactivity, this experimental results can be easily replicated even without interactivity (e.g. paper media). Thus, investigation of the influence of interactivity with different approaches from this study should be investigated.

Moreover, the social relationships among the participant, the PA, and the NPA were not understood by the participants until they filled in the blank thought balloon considering the situation among them. Consequently, understanding social relationships among the participant, the PA, and the NPA only with cover stories explained in the experiment and the initial interaction between the participant and PA seem to be hard for him/her. One of the solutions for these problems is to increase the opportunity for "rapport building" between a user and embodied agents [19]. Also considering the Moon's study [14] discussed in Section 3.3, it should be important to build rapport between a user and embodied agents whose perceived allocentric perspectives were different from the user's before the situation to enable a user to accept the agents' perceived allocentric perspective. Another solution should be addressing the procedural issues in the experiment, since the object-searching tasks were too hard for the participants to solve considering the situation among them, the PA, and the NPA. More understandable scenarios should be explored for the experiment.

6.3 Possible application and future work

In this experiment, we predicted that the user could consider the situation of his/her "enemy" through filling in a blank thought balloon emitted from the NPA and that changes would occur in the user's allocentric perspective, similar to changes that occur in the opinion of a debate participant after considering the thought of a "devil's advocate" [2]. As mentioned in Section 1, there are many situations in which people need to accept perspective from others in ordinary life. For example, in the situation that the user needs to try to take the minority's allocentric perspective to solve problems, an interface to let the user accept the minority's allocentric perspective should contribute to the user's solving of the problems. It is worth attempting to induce perspective sharing with the blank

thought balloon emitted from an embodied agent in other real problem solving situations.

7 Conclusion

In this study, through a psychological experiment that attempted to induce a user to accept perceived allocentric perspective of an embodied agent by filling in blank thought balloons, the possibility was explored of embodied agent technologies that let the user understand the perspective of others. The experimental results suggested that filling in a blank thought balloon emitted from the embodied agent may induce the user's acceptance of perceived allocentric perspective, but without establishment of social relationships between the user and the agent, the user has difficulty inspecting the perceived allocentric perspective. Finally, through introducing embodied agent technologies into practical situations, we intend to extract design principles of embodied agents that can let a user understand others' allocentric perspective that are different from his/hers.

References

[1] Norman, D.A.: Cognitive artifacts. In Carroll, J.M., ed.: Designing interaction: Psychology at the human-computer interface. Cambridge University Press, Cambridge (1991) 17–38

[2] Janis, I.L.: Groupthink: Psychological Studies of Policy Decisions and Fiascoes. 2nd edn. Houghton Mifflin Company, Boston, MA (1982)

[3] Morita, J., Miwa, K.: Changes of inferences caused by obtaining different perspectives: Analysis based on analogical reasoning. In: Proceedings of the 4th International Conference on Cognitive Science. (2003) 463–468

[4] Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco, CA (2003)

[5] Takeuchi, Y., Watanabe, K., Katagiri, Y.: Social identification of embodied interactive agent. In: Proceedings of the 13th International Workshop on Robot and Human Interactive Communication (RO-MAN 2004), Kurashiki, Japan (2004) 449–454

[6] Vogeley, K., Fink, G.R.: Neural correlates of the first-person-perspective. Trends in Cognitive Sciences **7**(1) (2003) 38–42

[7] Suzuki, S.V., Takeda, H.: Inducing change in user's perspective with the arrangement of body orientation of embodied agents. In: Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006). (in press)

[8] Okamoto, M., Okamoto, K., Nakano, Y.I., Nishida, T.: Supporting the creation of immersive CG contents with enhanced user involvement. In: Proceedings of the Symposium on Conversational Informatics for Supporting Social Intelligence and Interaction — Situational and Environmental Information Enforcing Involvement in Conversation, AISB'05: Social Intelligence and Interaction in Animals, Robots and Agents, Hatfield, UK (2005) 87–96

[9] Harrison, R.P.: The cartoon: Communication to the quick. Sage, Beverly Hills, CA (1981)

[10] Kurlander, D., Skelly, T., Salesin, D.: Comic Chat. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH96), ACM Press (1996) 225–236
[11] Sakamoto, R., Nakao, K., Sumi, Y., Mase, K.: ComicDiary: Representing individual experiences in comics style. In: Proceedings of the 25th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2001). (2001) 158
[12] Cooper, A., Reimann, R.M.: About Face 2.0: The Essentials of Interaction Design. Wiley, Indianapolis, IN (2003)
[13] Cassell, J., Sullivan, J., Prevost, S., Churchill, E., eds.: Embodied Conversational Agents. MIT Press, Cambridge, MA (2000)
[14] Moon, Y.: Intimate exchanges: Using computers to elicit self-disclosure from consumers. Journal of Consumer Research **26** (2000) 323–339
[15] Sundar, S.S., Nass, C.: Source orientation in human-computer interaction. Communication Research **27**(6) (2000) 683–703
[16] Wick, W.: Can You See What I See?: Picture Puzzles to Search and Solve. Cartwheel Books, New York (2003)
[17] Tajfel, H., Billig, M.G., Bundy, R.P., Flament, C.: Social categorization and intergroup behavior. European Journal of Social Psychology **1** (1971) 149–177
[18] Altman, I., Taylor, D.A.: Social penetration: the development of interpersonal relationships. Holt, Rinehart, and Winston, New York (1973)
[19] Bickmore, T., Cassell, J.: Small talk and conversational storytelling in embodied conversational interface agents. In: Proceedings of the AAAI Fall Symposium on Narrative Intelligence, Cape Cod, MA (1999) 87–92