Adaptation gap hypothesis: How differences between users' expected and perceived agent functions affect their subjective impression

Takanori Komatsu International Young Researcher Empowerment Center, Shinshu University Ueda, Nagano 3868567, Japan

and

Seiji Yamada National Institute of Informatics/SOKENDAI/Tokyo Institute of Technology Chiyoda, Tokyo 1018430, Japan

ABSTRACT

We describe an "adaptation gap" that indicates the differences between the functions of artificial agents that users expect before starting their interactions and the functions they perceive after their interactions. We investigated the effect of this adaptation gap on users' impressions of artificial agents because any variations in impression before and after the start of an interaction determines whether the user feels that this agent is worth interacting with. The results showed that positive or negative signs of the adaptation gap and subjective impression scores of agents before the experiment significantly affected the users' final impressions of the agents.

Keywords: human-agent interaction (HAI), adaptation gap, users' impressions of agents, gain and loss of esteem.

1. INTRODUCTION

Various interactive agents such as robotic agents [1] and embedded conversational agents (ECA) [2,3] have been developed to assist us with our daily tasks. In particular, researchers in the human-computer interaction and humanagent interaction communities are working hard to create such interactive agents. One of the hottest topics in these fields concerns how users' mental model of an agent affects their interaction with it. Because users supposedly base their mental model on the agent's appearance, its behaviors, and their preferences for the agent, the users' mental model significantly affects their interaction [4]. For example, when a user encounters a dog-like robot, s/he expects dog-like behavior from it, and s/he naturally speaks to it using commands and other utterances intended for real dogs, such as "sit," "lie down," and "fetch." However, s/he does not act this way toward a cat-like robot.

Several studies have focused on the effect of users' mental models about an agent on their interaction. Matsumoto et al. [5] proposed a "Minimal Design Policy" for designing interactive agents and concluded that the agent's appearance should be minimized in its use of anthropomorphic features so that users do not overestimate or underestimate the agents' competences. In fact, they applied this minimal design policy to developing Muu, their interactive robot [6] and Talking Eye, a life-like agent [7]. Kiesler [8] argued that agent design should include a process that anticipates a user's mental model about agents on the basis of the theory of common ground [9]; that is, individuals engaged in conversation must share knowledge (socalled, common ground) in order to be understood and have a meaningful conversation. Specifically, she stated that agents should be designed in such a way that a user could easily estimate her/his common ground (shared knowledge) with them. We believe that this design approach would be quite effective for users, especially at the beginning of an interaction, because it may determine whether or not a user would start an interaction with an agent.

2. ADAPTATION GAP BETWEEN A HUMAN AND AN AGENT

However, approaches like Matsumoto et al.'s [5] or Kiesler's [8] have a serious problem when an agent expresses behaviors that completely deviate from the users' mental model. Imagine that a user meets a human-like robot that looks very much like a real human being. This user would intuitively form a mental model of the robot, expecting human-like fluent speech and dexterous limb motions. However, if this particular robot could only express machine-like speech and halting limb motions that completely deviate from her/his mental model, s/he would be immediately disappointed with this robot because of its unexpected behaviors. The user would then stop interacting with it. To solve this problem, we need to determine the users' impressions of agents during their interactions; because such impressions would make users determine whether this agent is worth interacting with.

In this study, we focused on the difference between users' expectations regarding the function of the agent and users' actual perceived function, which is one of the factors that affects the users' impressions. We called this difference the adaptation gap (AG). Specifically, AG can be defined as $AG = F_{affer} - F_{before}$. Here, F_{after} is the function of the agent that a user actually perceived, and F_{before} is the users' expected function of the agent. We expect that this adaptation gap would have the following three properties [10,11].

- AG<0 (*F_{after}*< *F_{before}*): When users' expected function exceeds their perceived function, it is said that there is a negative adaptation gap. In this case, most people would be disappointed by the agent and would stop interacting with it.
- AG>0 (F_{after}> F_{before}): When users' perceived function exceeds their expected function, it is said that there is a positive adaptation gap. In this case, most people would not be disappointed by the agent and would continue interacting with it.
- AG=0 (F_{after} = F_{before}): When the perceived function equals the expected function, it is said that there is no adaptation gap. In this case, the agent would be regarded as just an instrument for users.

For example, when F_{before} is larger than F_{after} (say, when a user meets a human-like robot on the left side of Figure 1), AG would have a negative value (AG<0), and the user would most likely be disappointed. However, when F_{after} is larger than F_{before} (say, when a use meets a machine-like robot on the right side of Figure 1), AG would have a positive value (AG>0), and the user would be interested in interacting with this agent.



User's Mental Model of An Agent

Figure 2: Systematic concept of adaptation gap

The systematic concept of AG is depicted in Figure 2. Here, I_{before} is the users' subjective impressions about an agent before their interactions, and I_{after} is the users' impressions after the interaction. Note that these impressions (I_{before} and I_{after}) are totally different from the users' expected and perceived function of the agents (F_{before} and F_{after}); since F_{before} and F_{after} reflect the users' understanding of the behavioral aspects of the agent, e.g., whether the agent could achieve the given task or not, while I_{before} and I_{after} reflect the users' subjective impressions of the agent, e.g., whether this agent is a worthwhile interactive partner.

In particular, we expected that the value of AG strongly affects the user's impressions I_{after} . In this study, we then investigated the relationship between the values of AG and the user's subjective impressions after the interaction I_{after} (see the arrow on the right side of Figure. 2); that is, the independent variable in this concept is AG while the dependent variable is I_{after} . We assumed that this investigation would lead to verification of the above three properties about AG. Namely, if I_{after} is significantly influenced by AG (= F_{after} - F_{before}), we can conclude that the AG's properties have been verified.

3. EXPERIMENT

We conducted an experiment to investigate how the positive or negative signs of the adaptation gap affected the users' impressions of the agents.

Overview

We chose a "treasure hunting" video game as the experimental environment to observe the interaction between a user and an agent (Figure 3). In this game, a character on a computer monitor operated by a user walks on a straight road, with three tiny hills appearing along the way. A gold coin is inside one of the three hills, while the other two hills have nothing. The game ends after the character meets 100 sets of hills and the approximate duration of the game is about 10 minutes. The goal of the game is to get as many gold coins as possible.



Figure 3: Treasure hunting video game

In this experiment, the user got one point for each coin that s/he found during the first 80 trials, but from trials 81 on, s/he got ten points per coin. The participants were informed that 1 point was equivalent to 10 Japanese yen (about 10 US cents) and that, after the experiment, they could purchase some stationery (e.g., file holders or USB flash memories) of equivalent value with their points. The position of the coin in the three hills was randomly assigned. A robotic agent (explained later), which was placed next to the user, told the participant each time where it expected the coin would be. The participant could accept or reject the agents' suggestions.

Two different kinds of agents and participants

Twenty Japanese university students (10 men and 10 women; 21–23 years old) participated. These participants were randomly divided into the following two groups.

- AIBO Group (10 participants): The agent placed next to the participants was an AIBO robot (Sony Corporation, ERS-7, Figure 4 and 6). AIBO told the expected position of the coin by barking the number, e.g., one bark meant the first hill (the one on the left), two barks meant the middle hill, and three barks meant the third hill (the one on the right).
- MS Group (10 participants): The agent was a MindStorms (MS) robot (LEGO corporation, Figure 5 and 7). MS told the user the expected position by beeping the number, e.g., one beep meant the first hill, two beeps meant the second hill (middle), and three beeps meant the third hill.

The rate at which both agents succeeded in detecting the position of the coin was set at 60%, that is, the rates were equal. The behaviors of both agents were remotely operated by an experimenter in the next room performing in the Wizard of Oz (WOZ) manner. The treasure hunting video game was projected on a 46 inch LCD screen in front of the participants (Figure 8). Figures 6 and 7 show a photograph of the experiment.



Figure 4: AIBO robot (Sony Corporation, ERS-7)



Figure 5: MindStorms robot (LEGO corporation)



Figure 6: Photograph of AIBO group in experiment



Figure 7: Photograph of MS group in experiment



Analysis

We investigated the effect of the sign of the adaptation gap on the users' impressions of the agents. To acquire the users' impressions I_{before} and I_{after} , we had the participants fill in a questionnaire sheet based on the Liking scale [12], which consisted of 7 questions scored on a 9-point Likert scale. Accordingly, the summed score of these 7 questions (maximum: 63 points) was used as the participants' impression score for the agents (Table 1). A higher score indicated a better impression of the agents. In particular, before the experiment, the participants were asked to fill in this questionnaire about their expected functions of the agent F_{before} . After the experiment, they were asked to fill in the same questionnaire about their perceived functions of the agent F_{after} . Therefore, the signs of the adaptation gap were calculated by using the values of F_{before} and F_{after} ; that is, AG= $F_{after} - F_{before}$.

Table 1: Questionnaire given to acquire participants' impression scores of artificial agents before and after the experiment.

Questions

- 1 My feeling did not change even though this agent was next to me.
- 2 This agent has adequate adaptabilities for any user.
- 3 I can leave the important tasks to this agent.
- 4 This agent is quite good as a commercial product.
- 5 I have a high level of confidence in this agent.
- 6 This agent will become popular for most users.
- 7 This agent is the best product compared to the other kinds of agents.

Results

We assumed that the participants' impressions about the agents would be affected not only by the signs of the adaptation gap but also by the different agents or the achievement level of their interaction. Therefore, we investigated the effects of (1) the different agents (AIBO or MS), (2) the achievement level (high or low game score) and (3) the signs of the adaptation gap (positive or negative) on the users' impressions of the agents.

Case 1. Different agent: Figure 9 depicts the participants' impressions of the two agents, and the impressions scores of both groups after the experiment did not change much from before.



Figure 9: Impressions of agents of AIBO group and MS group

First, to see the effect of the differences in the agents on the participants' impressions, the participants' impressions in the AIBO and MS groups (10 participants each) were analyzed using a 2 (agent factor: AIBO or MS) x 2 (order factor: before or after the experiment) mixed ANOVA. The results showed no significant difference in interaction effects (F(1,18)=0.02, n.s.), and no main effects for agent or order factors (AIBO or MS: F(1,18)=0.32, n.s.; before or after: F(1,18)=0.18, n.s.).

Therefore, the differences in the agents did not affect the users' impressions of them.



Figure 10: Game scores of AIBO group and MS group

Next, to see the effect of the different agents on the achievement level, the game scores in both groups were analyzed using a 2 (AIBO or MS) ANOVA. No significant difference in game scores (AIBO: 140.6 points, MS: 143.8 points, F(1,18)=0.10, n.s.) was obtained (Figure 10).



Figure 11: Expected and perceived functions of AIBO group and MS group

Finally, the expected and perceived functions of the agents (F_{before} and F_{after}) were analyzed using a 2 (AIBO or MS) x 2 (before or after) mixed ANOVA (Figure 11). Again, no significant differences were obtained for interaction effects (F(1,18)=0.21, n.s.) or both main effects (AIBO or MS: F(1,18)=0.23, n.s.; before or after: F(1,18)=0.01, n.s.). The average value of the adaptation gap in the AIBO group was -1.9 (S.D.=20.5), and the average for the MS group was 4.9 (S.D.=23.8). Therefore, the differences between the agents did not affect the achievement level of the interaction or the expected and perceived functions.

To sum up, the differences in the agents did not affect the users' impressions of the agents, the achievement level of interaction, or the expected and perceived functions.

Case 2. Achievement level (Game score): All 20 participants, regardless of whether they were in the AIBO or MS group, were divided into two groups in terms of acquired game scores: high scoring and low scoring (10 participants in

each group). The average game score of the high scoring group was 158.5 points, while the average score of the low scoring group was 125.9 points. That is, there was a significant difference in average game score between high and low scoring groups (F(1,18)=26.3, p<.01(**), Figure 12).



Figure 12: Game scores of high-scoring group and low-scoring group



Figure 13: Impressions of agents of high-scoring group and low-scoring group

Figure 13 depicts the participants' impressions of the agents. The impression scores of both groups after the experiment did not change from before. The participants' impressions of the agents were then analyzed using a 2 (score factor: high or low) x 2 (order factor: before or after) mixed ANOVA. The results showed no significant differences in interaction effects (F(1,18)=0.04, n.s.) between groups or on each of the main effects (high or low: F(1,18)=1.06, n.s.; before or after: F(1,18)=0.18, n.s.). Therefore, game score as an achievement level of interaction did not affect the users' impressions of the agents.

We also analyzed the expected and perceived functions of the agents (F_{before} and F_{after}) using a 2 (high or low) x 2 (before or after) mixed ANOVA (Figure 14), and the results showed no significant differences in interaction effects (F(1,18)=0.03, n.s.) or in both main effects (high or low: F(1,18)=2.84, n.s.; before or after: F(1,18)=0.01, n.s.). The average value of the adaptation gap in the high group was 1.6 (S.D.=23.4), and the average value in the low group was -0.5 (S.D.=21.1).



Figure 14: Expected and perceived function of high-scoring and low-scoring group

To sum up, the achievement level of the game (game score) did not affect the users' impressions of the agents or their expected and perceived functions.

Case 3. Signs of adaptation gap: All 20 participants, regardless of whether they were in the AIBO or MS group, were divided into two groups based on the signs of the adaptation gap ($AG=F_{after}-F_{before}$): positive (AG>0; $F_{after}>F_{before}$, 9 participants) or negative (AG <0; $F_{after}<F_{before}$, 10 participants) groups. Note that the one participant who had an AG of 0 was eliminated from this analysis. The average expected functions of the agents, F_{before} , were respectively 42.6% and 68.7% in the positive and negative groups, while the average perceived functions, F_{after} , were 63.4% and 51.0%. Thus, the average value of the adaptation gap in the negative group was -17.7 (S.D.=17.7), and it was 20.8 (S.D.=11.8) in the positive group (Figure 15).



Figure 15: Expected and perceived function of positive group and negative group.

Figure 16 depicts the participants' impressions of the agents. The impression scores in the positive group increased, while the ones in the negative group decreased. The participants' impressions of the agents in the positive and negative groups were then analyzed using a 2 (positive or negative) x 2 (before or after) mixed ANOVA. The results showed a significant difference in the interaction effects (F(1,17)=14.29, p<.01 (**)), and there were significant differences in the participants' impressions before and after the experiment in both groups (positive: F(1,17)=9.94, p<.01 (**); negative: F(1,17)=4.81,

p<.05 (*)) and a significant tendency in the impressions before the experiment between the two groups (F(1,17)=3.99, p<.1 (+)). Therefore, the signs of the adaptation gap actually affected the users' impressions of the agents; specifically, the impressions in the positive group significantly increased (about 4.9 points), while the ones in the negative group significantly decreased (about 3.4 points).



Figure 16: Impressions of agents of positive group and negative group.

However, from Figure 16, it seemed that the participants' impression scores acquired before the experiment also had some effect on the final users' impression scores after the experiment. To confirm these effects, we performed a 2 (positive or negative) ANCOVA (independent variable: positive or negative groups, covariate: impression scores before the experiment, dependent variable: impression scores after the experiment), and the results showed significant differences in the main effect of the independent variable (F(1,16)=11.69, p<.01 (**)) and in the main effect of the covariate (F(1,16)=28.89, p<.01 (**)). Therefore, the final impression scores of the agents acquired after the experiment were affected by two factors simultaneously; one was the signs of the adaptation gap (independent variable), and the other was the impressions scores before the experiment (covariate).



Figure 17: Game scores of positive group and negative group.

While no significant differences were found in the game score between the two groups (positive: 145.9, negative: 138.8; F(1,17)=0.46, n.s.), we found that the signs of the adaptation gap did not affect the achievement level of interaction (Figure. 17).

To sum up this analysis, the signs of the adaptation gap did not affect the achievement level of the interaction but the participants' impressions of the agents. Moreover, the impression scores before the experiment also affected the variations of their impressions.

4. DISCUSSION

Summary and future plans

The results of our experiment can be summarized as follows:

- The differences in the agents (i.e., whether the agent was AIBO or MS) did not affect the users' impressions of the agents.
- The achievement level of the interaction (i.e., whether the participants could get the higher scores or not) did not affect the users' impressions.
- Two factors, signs of the adaptation gap and impressions scores before the experiment, significantly affected the users' impressions of the agents. Specifically, the participants, who showed positive signs of the adaptation gap and lower impressions scores of the agents before the experiment, ended up having higher impressions of the agents, whereas the participants, who showed negative signs of the adaptation gap and higher impressions scores before the experiment, ended up having lower impressions of the agents.

Therefore, not only the signs of the adaptation gap but also users' impressions before the experiment had a significant effect on the users' final impression of the agents. This means that the users' impressions would not become higher even if the adaptation gap was positive. Therefore, these results suggest a rigid condition regarding the adaptation gap; one must comprehend the users' impressions of the agents before the interaction in order to have the properties of the adaptation gap described in section 2 (especially in the case of "AG<0" and "AG>0").

One of the interesting phenomena of this experiment was that there were no significant differences in the expected function of the agents F_{before} and in the impressions scores of the agents before the experiment I_{before} between the AIBO and MS groups. We assumed that this phenomenon was due to the fact that the agents' appearances do not have a strong enough effect to make the participants evoke uniform F_{before} and I_{before} . This means that the participants' mental models about the agents were diverse. Therefore, Kiesler's argument of [5] that agents should include a process that anticipates a user's mental model about agents should be carefully applied, considering not only the agents' appearance but also its given tasks, relationship with users, and so on.

However, it can be said that the relationship between the independent variable (signs of adaptation gap) and dependent variable (participants' final subjective impression) was a bit complicated in this experiment. In order to simplify this relationship, we are planning to construct a new experimental setting (Figure 18). Specifically, this involves preparing two experiments; the first experiment is to measure the sign of the adaptation gap, and the second experiment is to measure participants' subjective impressions. After the first experiment, the participants are divided into two experimental groups, an "AG >0 group" and an "AG <0 group", based on the signs of AG. After the second experiment, the subjective impression

scores of participants are analyzed by comparing these two groups. We believe that this experimental plan would clarify and refine the concept of the adaptation gap hypothesis.



Figure 18. Plan of new experiment to clarify the relationship between independent variable and dependent variable

After resolving the above issue, we will conduct a subsequent study to develop a practical design methodology on how to make users evoke uniform F_{before} and impressions before interaction I_{before} . If we succeed in developing such a methodology and set up a lower F_{before} and lower I_{before} we could create a "positive adaptation gap" situation in which the participants end up with higher impressions of the agents by which F_{after} is higher than F_{before} . Such a consecutive study would be the basis of a novel design methodology to make agents that users would want to continue interacting with and without losing their good impressions, and such a methodology would make it possible for various interactive agents with which users would find worth interacting.

Relevance to existing findings

The results presented here are quite similar to findings in social psychology on "gain and loss of esteem" [13]; that is, person A prefers person B who changed her/his evaluation of person A from positive to negative, to person C who maintained a positive evaluation (gain effect), and person A has a worse perception of person B who changed her/his positive evaluation into a negative one, compared with person C who maintained a negative evaluation (loss effect). This implies that people's impressions about others are determined not only by the total evaluation given by others but also by the latest evaluation when it changes significantly. While the gain and loss of esteem states that people's impressions are determined by the objective evaluations received from others, our results show that users' impressions of agents are determined by the subjective evaluations they themselves generated. Therefore, the results of this study did not simply replicate the existing findings in HCI field on media equation studies [14], etc.; it generated new findings about significant personal properties that are applicable to various HCI areas.

Moreover, the acquired results have bearing on the "uncanny valley" [15]; that is, agents are becoming similar in appearance to human beings. At some point, people will start feeling uncanny or loss of familiarity when dealing with such agents because small differences in appearances of the agents from those of humans would become extremely emphasized. Figure 19 qualitatively illustrates this relationship between the agents' likeness to human beings and the familiarity that human users feel.



Figure 19. Concept diagram of uncanny valley¹

We believe that the reason why humans suddenly start feeling uncanny can be explained by our adaptation gap hypothesis, because the "human-likeness" of the agents would form an expectation about them in the user's mind, and "familiarity" would be determined by the perceived functions of the agents. Therefore, we could assume that the relationship between human-likeness and familiarity of the agents can be replaced by one between the expected function and perceived functions of the agents. If we could measure the familiarity of the agent as a dependent variable in our experiment, the above assumption might be able to be verified, and the results might contribute to development of a novel methodology for designing interactive agents.

5. CONCLUSIONS

We described an "adaptation gap" that indicates the differences between the functions of artificial agents users expect before starting their interactions and the functions they perceive after the interactions. We investigated the effects of this adaptation gap on users' impressions of the artificial agents because any variations in impressions before and after the start of an interaction determine whether the user feels that this agent is worth interacting with. The results showed that the positive or negative signs of the adaptation gap and the subjective impression scores of the agents before the experiment significantly affected the final users' impressions of the agents.

Thus, this study could contribute to the development of a novel methodology to make users continue interacting with agents without losing their impressions of them, and such a methodology could be easily applied to various interactive agents that could make users believe is worth continuing an interaction. Although these results were acquired in an organized experimental environment and their applicability to actual usage is still unclear, a comparison with other findings, such as on "gain and loss of esteem" or the "uncanny valley hypothesis", would contribute to the development of a design methodology for interactive agents.

This figure is copied from

http://en.wikipedia.org/wiki/Uncanny_valley

ACKNOWLEDGEMENTS

This study was partially supported by Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

REFERENCES

- M. Imai, T. Ono, & H. Ishiguro, "Robovie: Communication Technologies for a Social Robot." International Journal of Artificial Life and Robotics, 2003, pp. 73-77.
- [2] H. Prendinger, & M. Ishizuka, Life-Like Characters. Springer, 2004.
- [3] J. Cassell, T. Stocky, T. Bickmore, Y. Gao, Y. Nakano, K. Ryokai, D. Tversky, C. Vaucelle, & H. Vilhjalmsson, "MACK: Media lab Autonomous Conversational Kiosk." Proceedings of Imagina02, 2002.
- [4] T. Komatsu, & S. Yamada, "How do robotics agents' appearances affect people's interpretations of the agents' attitudes?" Extended Abstract of ACM-CHI2007 (workin-progress), 2007, pp. 2519-2525.
- [5] N. Matsumoto, H. Fujii, M. Goan, & M. Okada, "Minimal design strategy for embodied communication agents." Proceedings of the 14th International Symposium on Robot and Human Interactive Communication (RO-MAN2005), 2005, pp. 335-340.
- [6] M. Okada, S. Sakamoto, & N. Suzuki, "Muu: Artificial Creatures as an Embodied Interface." Proceedings of 27th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2000), 2000, p. 91.
- [7] N. Suzuki, Y. Takeuchi, K. Ishii, & M. Okada, "Talking Eye: Autonomous Creatures for Augmented Chatting." Robotics and Autonomous Systems, Vol. 31, 2000, pp. 171-184.
- [8] S. Kiesler, "Fostering common ground in human-robot interaction." Proceedings of the 14th International Symposium on Robot and Human Interactive Communication (RO-MAN2005), 2005, pp. 729-734.
- [9] H. H. Clark, Arenas of language use, University of Chicago Press, Chicago, USA, 1992.
- [10] S. Yamada, & T. Komatsu, "Designing Simple and Effective Expressions of Robot's Primitive Minds to a Human." Proceedings of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2006), 2006, pp. 2614-2619.
- [11] T. Komatsu, & S. Yamada, "Effect of agent appearance on people's interpretation of agent's attitudes." Extended Abstract of ACM-CHI-2008 (work-in-progress), 2008, pp. 2929-2934.
- [12] Z. Rubin, "Measurement of romantic love." Journal of Personality and Social Psychology, Vol. 16, 1970, pp. 265-273.
- [13] R. Aronson, & D. Linder, Gain and loss of esteem as determinants of interpersonal attractiveness. Journal of Experimental Social Psychology, Vol. 1, 1965, pp. 156-171.
- [14] B. Reeves, & C. Nass, **The Media Equations**, CSLI Publications, USA, 1998.
- [15] M. Mori, "Bukimi no tani (The Uncanny Valley)." Energy, Vol. 7, No. 4, 1970, pp. 33–35 (Originally in Japanese)