User Reactions to Anthropomorphized Interfaces

Tomoko Koda

JST CREST Digital City Project/Kyoto University, Kyoto 604-0924 Japan

Abstract

It is still an open question whether software agents should be personified in the interface. In order to study the effects of faces and facial expressions in the interface, a series of experiments was conducted to compare subjects' responses to and evaluation of different faces and facial expressions.

The experimental results obtained demonstrate that: 1) personified interfaces help users engage in a task, and are well suited for an entertainment domain; 2) people's impressions of a face in a task are different from ones of the face in isolation. Perceived intelligence of a face is determined not by the agent's appearance but by its competence; 3) there is a dichotomy between user groups which have opposite opinions about personification. Thus, agent-based interfaces should be flexible to support the diversity of users' preferences and the nature of tasks.

Keywords

agent, anthropomorphism, face, facial expression, user interface design, individual differences, emotion

INTRODUCTION

There is a growing interest in software agents which assist a user in performing daily computer tasks. Software agents will make software more active and work autonomously without waiting for a user's command. Negroponte suggests that this "agent-based" interface should be social and active to support cooperative tasks between a user and an agent [10]. Maes addresses that the key HCI issue in building a successful software agent is to help the user understands what the agent's capacities are, its limitations, and its way of operating, etc. [8].

Many software agents are personified in terms of a human or caricature face. "Phil" in Apple's Knowledge Navigator video [7] and Apple's "Guides" project [11] used salient human figures as interface metaphors. The work on social interface by Nass et al. [9] has been used as a basis of several studies, such as Microsoft's Persona Project and their software products [2]. These studies hypothesize that if the interaction with computers has a social nature, using a character in the interface is a natural way to support the interaction. The MIT Media Lab's software agents, such as the Maxims system [6] and Newt [13], used faces to help users understand the agent's characteristics or behaviors. This research hypothesizes that users interact with agents more comfortably when they are able to predict the agent's characteristics and behaviors on the basis of its external traits. Though the research described so far reports positive results for a personified interface, the results are obtained by observation or interviews, not by quantitative analysis. The primary focus of their research is on implementation of a prototype rather than on user testing.

However, other research with quantitative analysis indicates that adding a face to an interface does not necessarily result in better human-computer interactions. Walker et al. [15] report that having a face is engaging and takes more effort and attention from the user. A face with more expression in an interface leads to greater engagement, but does not always makes the experience for users if it is added incautiously. Takeuchi et al. [14] report that users respond differently to systems having a face than to those without. They suggest that a face in an interface takes more effort from the user because people try to interpret the human images.

The research reported in this paper includes both implementation and quantitative analysis of subjects' impression about a personified interface. This paper argues that employing a face as the representation of an agent is engaging and makes a user pay attention. As Don [3] says, personification can be useful if we understand its strengths and weaknesses. It is possible to make advantages of personification outweigh the drawbacks by using faces in appropriate application domains, where conveying the agent's behavior to the user is important for making a user comfortable with the system, or in an entertainment domain where engagement is crucial and users don't try to minimize efforts. This paper reports the results of an investigation on people's impressions of faces in an interactive entertainment environment.

DESIGN OF THE EXPERIMENTS

Web-based Experiments

We built an application to test agent-based interfaces, namely a poker game, and experiments were performed on the World Wide Web (see [5] for the URL of the poker game) so as to allow users to participate in the experiment exactly the same way as they work in normal conditions. There is a need to conduct experiments on the World Wide Web because of two reasons. One is that the nature of the experiment needs participants from various backgrounds and age ranges. The other is that current software agents are mostly used on the World Wide Web.

Research Questions

The questions asked in this study are as follows:

1) Do people pay attention to a face or facial expressions of an agent?

2) Are people distracted by the face or facial expressions?

3) Does having a face increase people's engagement in a task?

4) Do people use information from the face to interpret the agent's behavior?

5) What kind of facial features (gender, level of humanity, level of realism) make the agent look intelligent, likable, and comfortable to interact with?

6) Is people's impression of an agent determined based on its appearance or on its performance or both?

7) Does people's opinion about personification affect their impression of the faces?

8) Is there any difference between the subject's gender in subjects' impression of faces?

System Design

We designed and built a web-based poker game in which computer poker playing agents play against each other and a user. The task for the user is to play this poker game against the poker playing agents which have personified representations. The game of poker was chosen to observe to what extent subjects concentrate just on their hands, or pay attention to the faces of opponents and try to interpret their facial expressions. Figure 1 shows the interface for the game.

Figure 1: Web interface for the poker game (HUMANITY experiment)

Character Design

The poker playing agents used in the experiments have seven different graphical representations: Realistic Human Male/Female faces, Caricature Human Male/Female faces, a Caricature Dog's face, a line-drawn Smiley face, and the Invisible Man, which is a white square without a face (see Figure 2). They vary in terms of gender, humanity, and realism.

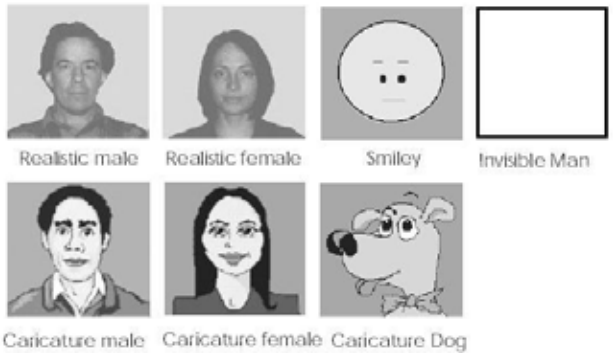

Figure 2: The graphical representations of the poker playing agents

Facial Expression Design

Each agent has ten facial expressions (neutral, pleased, displeased, satisfied, disappointed, surprised, relieved, excited, anxious, and very excited). It changes its expressions when it deals, bets, and wins/loses. For example, the Dog's face might show an exited expression while betting, or show a disappointed face when it loses. Figure 3 shows the ten expressions of the Caricature Female face in the same order as described above.

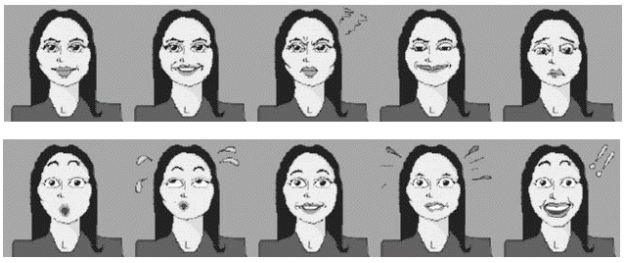

Figure 3: Ten facial expressions of the Caricature Female face.

The emotions were derived from the Ortony, Clore, and Collins (OCC) model of emotion types [12]. The OCC model assumes that emotions are the result of valanced--positive or negative--reactions to situations which a person experiences. In the poker game situation, we assumed that the world consists of events--i.e. having good hand or raising $10--, which lead to the ten emotions described above. Figure 4 shows the rules that lead to the ten emotion types which agents express while playing the poker game.

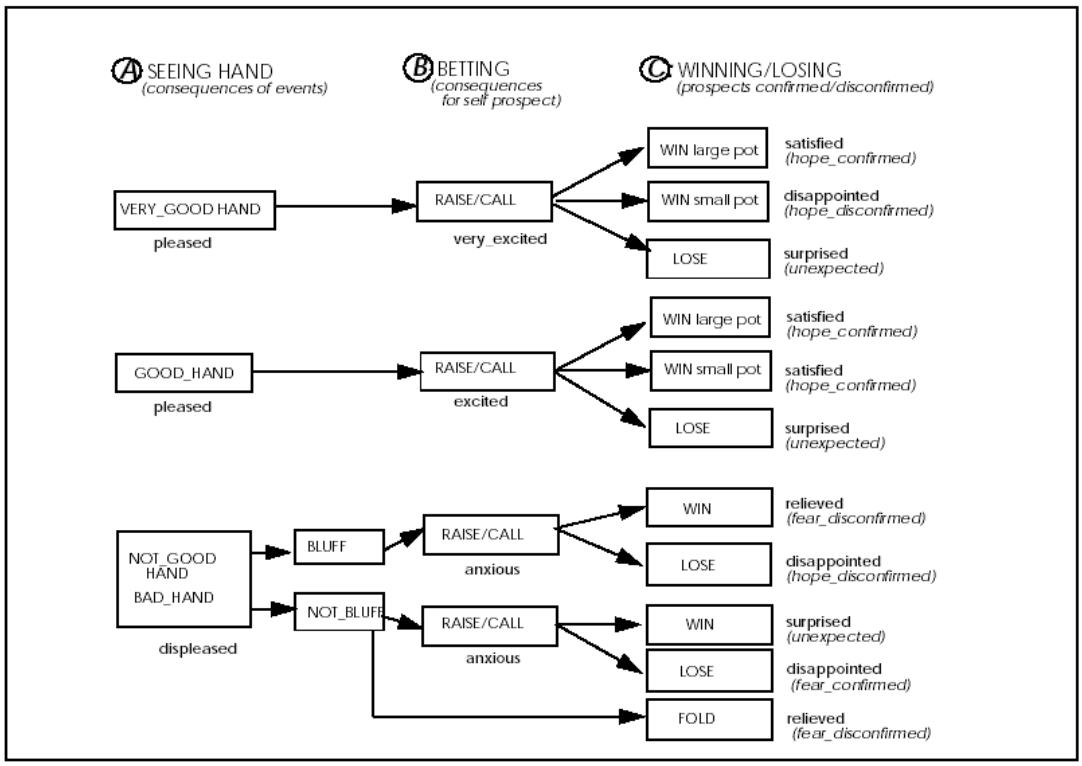

Figure 4: The rules that lead to the ten emotion types expressed by the agents while playing the poker game

Procedure

A typical scenario of an experiment is as follows:

A subject accesses the poker game web site [5]. The subject reads the consent form and instructions, agrees to participate in the experiment, and submits an electronic agreement form. The subject sees the poker playing agents faces to familiarize himself/herself with the agents' faces and facial expressions. The poker game program is downloaded to the subject's computer from the server computer. The subject plays a poker game with one out of five conditions for 15 rounds, which takes from 10 to 15 minutes to complete. The subject's and agents' hands, actions, and the ending money state are logged into the server. A questionnaire appears after playing 15 rounds. The subject answers and submits the questionnaire. The answers are stored in a server disk.

Conditions of the Experiments

There are five web-based experiments which are randomly assigned to a user. The FACE experiment is designed to study the effect of having a face. The representations are the Caricature Female face and the Invisible Man. The GENDER experiment studies the difference in subjects' impressions of characters' gender. The representations used in this experiment are the Realistic Male and the Realistic Female face. The HUMANITY experiment is designed to compare the difference in subjects' impressions between a human face and a non-human face. The representations used are the Caricature Male face and the Dog's face. The REALISM experiment uses three faces to compare three different levels of realism. The representations used are the Realistic Male face, the Caricature Male face, and the Smiley face. The EXPRESSIVENESS experiment has three players with the same facial representation. This experiment was designed to evaluate the difference in subjects' impression between three modes of expressiveness. The Caricature Female face with different color of clothes is used to distinguish each face. One agent has Honest expressions, the other has Deceiving expressions, i.e. the player is expressive but the facial expressions don't correspond to their actual emotions, and the last one has Stoic expressions, i.e. the player shows only neutral and pleased expressions regardless of its hand. Table 1 shows the facial representations used in each experiment. All facial images used in the experiments have the same size, the same resolution, and the same background color. Except for the EXPRESSIVENESS experiment, all poker playing agents are in Honest mode. The agents play poker game using exactly the same poker playing strategy, which the subjects were not made aware of.

Table 1: The facial representations used in each experiment

	Player 1	Player 2	Player 3
FACE	Caricature Female	Invisible man	-
GENDER	Realistic Male	Realistic Female	-
HUMANITY	Caricature Male	Caricature Dog	-
REALISM	Realistic Male	Caricature Male	Smiley
EXPRESSIVENESS	Honest Caricature Female	Deceiving Caricature Female	Stoic Caricature Female

Lab Experiments

Three pilot experiments were conducted prior to the five web-based experiments. The purpose of the pilot experiments was to collect subjects' impressions of each face based solely upon its visual appearance. Subjects were shown one of the sets of faces used in the GENDER, HUMANITY, and REALISM experiments with a neutral expression. The subjects answered questions about each face's perceived intelligence, likability, and engagingness as an opponent poker player, just by looking at the facial representations.

In addition to the web-based experiments, seven subjects participated in a lab experiment. The lab experiment used exactly the same procedure as the web-based ones, except that subjects were videotaped while playing the poker game and had interviews about their experience after playing the poker game.

Questionnaire

Questions varied according to the experiment. There are 4 background questions (subjects's gender, age, computer expertise, and their opinion about personification), and 4 to 8 questions about the subjects' impressions of the faces against which they played poker. The questions about their impressions are answered on a seven point scale (1: strongly disagree to 7: strongly agree).

The common questions across the experiments are perceived intelligence, likability, engagingness, and level of comfort for each face. Perceived intelligence is measured by how good at playing poker each player is perceived to be. Likability is measured by how much the user say they enjoyed playing against each player. Engagingness is measured by how engaging playing poker against each player is considered to be. Comfortableness is measured by how much the user wants to play poker with the same player.

Specific questions are required attention and level of distraction (for the FACE and the EXPRESSIVENESS experiments), level of correspondence to the actual intelligence of the player (for the HUMANITY and the REALISM experiments), believability of the facial expressions, and the usefulness of the facial expressions (for the EXPRESSIVENSS experiment). To measure level of distraction, subjects are asked to rate how much they were distracted by the existence of the face or facial expressions on a 7 point scale. Subjects measure the level of correspondence of a face by how suitable the face is in representing the actual poker playing skill of the player. Believability is measured by how much the user believed each player was honest in its facial expressions about its poker hand. Subjects measure the usefulness of a player's facial expressions by how helpful those facial expressions are in understanding the player's strategy.

Subjects

Experiments continued for one and a half months. Subjects participated in the experiments voluntarily from all over the world using the World Wide Web. More than 1,000 people accessed the poker game site and 157 of them answered the questionnaire, for a response rate of 15%. Of these subjects, 78% were male. The age range of the subjects was from 10 to 50 years old. Fifty-seven percent of them were in their 20's, 26% were teens, and 14% were in their 30's. Fifty-two percent of them rated themselves as advanced computer users, 40% as intermediate users. When asked about personifying an interface, 51% of them supported having a face on the screen (AGREE group), the rest were against having a face (DISAGREE group).

RESULTS

Impressions Based on Appearance

The pilot experiments evaluate subjects' impressions based solely on agents' appearance, in particular, gender, humanity, and realism. The result of the comparison for gender shows that there is no difference between people's impression of the Male face and Female face in terms of its perceived intelligence, likability, and engagingness, as shown in Table 2. The result of the difference for level of humanity shows that the Human face is perceived as more intelligent than the Dog's face based on visual appearances ($t(15)=3.10$, $p<.01$), but less likable ($t(15)=-3.13$, $p<.01$) and engaging ($t(15)=-5.57$, $p<.01$) as a representation for a poker player, as shown in Table 3. In the experiment to compare different levels of realism, subjects rated the Realistic face to be more intelligent ($F(6.12,2)$, $p<.01$), likable ($F(5.06, 2)$, $p<.05$), and engaging ($F(6.25, 2)$, $p<.01$) than the Caricature face and Smiley, as shown in Table 4.

Table 2: The mean value of impressions for the Male and Female face based on appearance

Variables	Male (n=15)	Female (n=15)	t (paired)
INT (Intelligence)	4.73	5.13	t(14) = -.88
LIKE (Likability)	4.80	4.73	t(14) =.19
ENG (Engagingness)	5.13	4.80	t(14) =.77

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

Table 3: The mean value of impressions for the Human and Dog's face based on appearance

Variables	Human (n=16)	Dog (n=16)	t (paired)
INT (Intelligence)	3.88	2.63	t(15) = 3.10*
LIKE (Likability)	3.81	5.37	t(15) = -3.13*
ENG (Engagingness)	3.50	5.56	t(15) = -5.57*

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

Table 4: The mean value of impressions for Smiley, the Caricature, and the Realistic face based on appearance

Variables	Smiley(n=16)	Caricature (n=16)	Realistic (n=16)	F(value, df)
INT (Intelligence)	3.53	3.59	4.94	F (6.12, 2)*
LIKE (Likability)	3.76	4.06	4.88	F (5.06. 2)**
ENG (Engagingness)	3.65	3.82	5.12	F (6.25, 2)*

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

Difference for Face and Noface

As shown in Table 5, subjects paid attention to the face (ATT=5.28, where 4.0 is neutral), but were not distracted by the existence of the face (DIS=2.84, where 4.0 is neutral). Observations and interviews in the lab experiment support this result. Subjects rated having a face to be significantly more likable (t(24)=2.76, p<.05, see Figure 5), more engaging (t(24)=2.85, p<.01), and more comfortable to play against (t(24)=3.36, p<.01), than not having a face. However, both poker playing agents are rated to be equally intelligent regardless of the existence of a face.

Table 5: The mean value of impressions for the Caricature face (Face) and the Invisible Man (NoFace) in the game

Variables	Face (n=25)	NoFace (n=25)	t (paired)
INT (Intelligence)	3.64	3.80	t(24) = -.27
LIKE (Likability)	5.08	3.72	t(24)=2.76**
ENG (Engagingness)	4.60	3.32	t(24) = 2.85*
COM (Comfortableness)	4.88	3.36	t(24) = 3.06*
ATT (Attention to the face)	5.28	-	-
DIS(Distracted by the face)	2.84	-	-

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

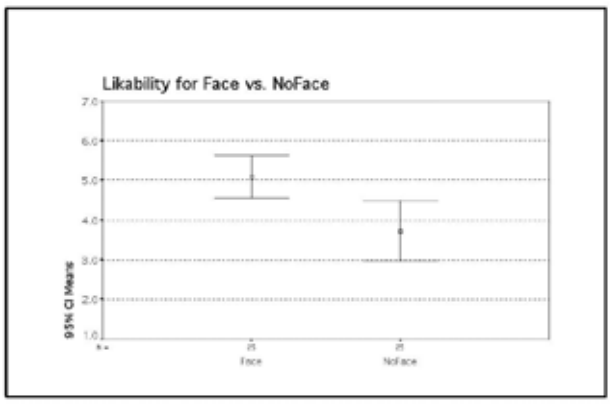

Figure 5: The mean value of likability for the Face vs. NoFace condition

Difference for Characters' Gender

As shown in Table 6, there is no main effect of the agent's gender in any variables. However, there are significant interactions between people's opinion about personification (VOTE) and perceived intelligence (F(4.73, 1), p<.05), and slight interactions between VOTE and likability (F(3.62, 1), p<.10). The AGREE group (those who are for personification) rated perceived intelligence and likability for the Male face higher and the Female face lower. While the DISAGREE group (those who are against personification) rated the Male and the Female face in the opposite way. Figure 6 shows the mean value of perceived intelligence of each face rated by each group. No other 2-way or 3-way interactions were found.

Table 6: The mean value of impressions for the Male and Female face in the game

Variables	Male (n=37)	Female (n=37)	t (paired) F(value, df)
INT (Intelligence) Vote Agree (n=18) Vote Disagree (n=19)	3.35 3.94 2.78	3.83 3.55 4.10	t(36) = -1.33 F(4.73, 1)**
LIKE (Likability) Vote Agree (n = 18) Vote Disagree (n = 19)	4.49 5.11 3.89	4.62 4.83 4.42	t(36) = -.55 F(3.62,1)***
ENG (Engagingness)	4.32	4.43	t(36) = -.55
COM (Comfortableness)	4.00	3.92	t(36) =.30

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

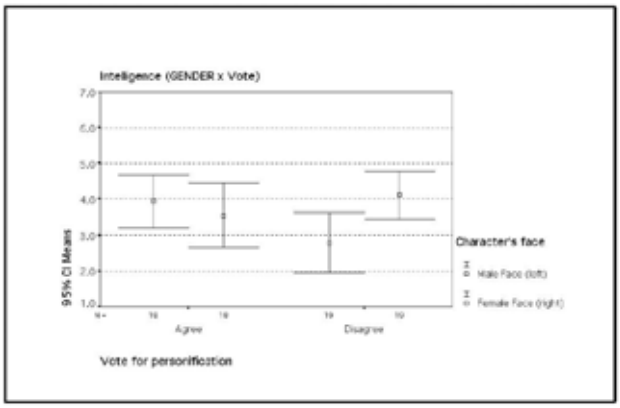

Figure 6: The mean value of perceived intelligence for the Male and Female face rated by the AGREE and DISAGREE group

Difference for Level of Humanity

As shown in Table 7, there is no main effect of level of humanity (Human or Dog's face) in any variables. However, there are significant interactions between VOTE and likability (F(5.40, 1), p<.05), comfortableness (F(11.42, 1), p <.01), and level of correspondence to the actual poker playing skill (F(6.97, 1), p <.05). The AGREE group rated likability, comfortableness, and correspondence of the Human face significantly lower,

while they rated the Dog's face significantly higher. The DISAGREE group rated the Human and the Dog's face in the opposite way. Figure 7 shows the mean value of comfortableness of each face rated by each group. There are also interactions between subjects' gender and likability (F(9.24, 1), p<.01), engagingness (F(5.94, 1), p<.05), and comfortableness (F(7.26, 1), p<.05). The result suggests that male subjects rated the Dog's face as more likable, engaging, and comfortable than the Human face, while female subjects rated in the opposite way.

Table 7: The mean value of impressions for the Human and Dog's face in the game

Variables	Human (n=23)	Dog (n=23)	t (paired) F(value, df)
INT (Intelligence)	3.73	3.82	t(22) = -.18
LIKE (Likability)	4.56	4.39	t(22)=.53
Vote Agree (n = 11)	4.72	4.73	F(5.40,1)**
Vote Disagree (n = 12)	4.45	4.09	
Male subject (n = 17)	4.47	4.65	F(9.24,1)*
Female subject (n = 6)	4.83	3.67	
ENG (Engagingness)	3.95	4.13	t(22) = -1.00
Male subject (n = 17)	3.88	4.29	F(5.94, 1)**
Female subject (n = 6)	4.17	3.67	
COM (Comfortableness)	4.08	4.04	t(22) =.13
Vote Agree (n = 11)	3.91	4.55	F(11.42, 1)*
Vote Disagree (n = 12)	4.27	3.55	
Male subject (n = 17)	4.29	4.41	F(7.26, 1)**
Female subject (n = 6)	3.50	3.00	
COR (Correspondence)	3.91	3.50	t(22) = 1.16
Vote Agree (n = 11)	3.82	4.18	F(6.97, 1)**
Vote Disagree (n = 12)	4.00	2.82	

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

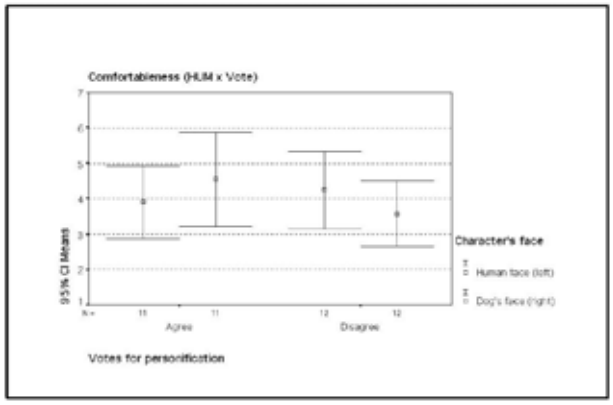

Figure 7: The mean value of comfortableness for the Human and Dog's face rated by the AGREE and DISAGREE group

Difference for Level of Realism

As shown in Table 8, there is no main effect of level of realism in perceived intelligence, likability, and engagingness. The results show that the Realistic face is slightly more intelligent, likable, engaging than the Caricature and Smiley, though not significantly so. There are main effects of level of realism in comfortableness (F(.05, 2), p<.10, see Figure 8) and correspondence to the skill level (F(.18, 2), p<.5, see Figure 9). The Realistic face is rated as more comfortable to play against than the Smiley face. While in terms of correspondence to the actual poker playing skill, subjects rated the Smiley face as the most appropriate representation, and the caricature and Realistic faces are less appropriate. No 2-way or 3-way interactions were found.

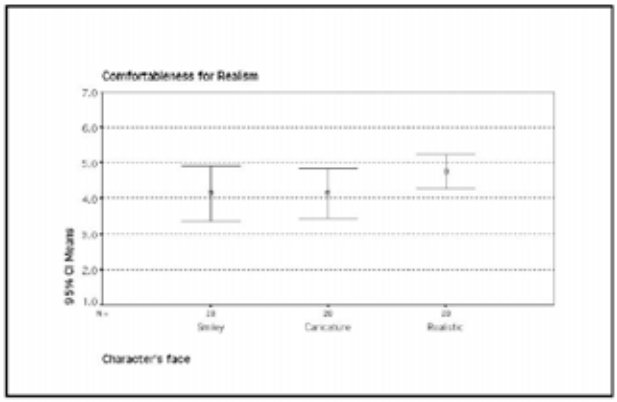

Figure 8: The mean value of comfortableness for Smiley, the Caricature, and the Realistic face

Table 8: The mean value of variables for Smiley. Caricature, and Realistic face in the game

Variables	Smiley (n=30)	Caricature (n=30)	Realistic (n=30)	F(value, df) t (paired)
INT (Intelligence)	3.37	3.53	3.70	F (.30, 2)
LIKE (Likability)	4.20	4.10	4.70	F (.59, 2)
ENG (Engagingness)	4.17	4.13	4.76	F (.58, 2)
COM (Comfortableness)	4.07	4.00	4.63	F (.05, 2)***
Smiley vs. Caricature				t(29) =.14
Caricature vs. Realistic				t(29) = -1.60
Smiley vs. Realistic				t(29) = -1.71***
COR (Correspondence)	4.67	3.90	4.03	F (.18, 2)**
Smiley vs. Caricature				t(29) = 2.25**
Caricature vs. Realistic				t(29)= -.36
Smiley vs. Realistic				t(29) = 2.00**

** p<.01, ** p<.05, *** p<.10, no mark = not significantly*

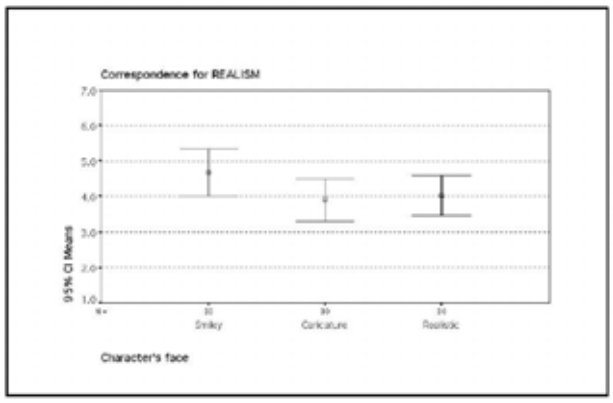

Figure 9: The mean value of correspondence for Smiley, the Caricature, and the Realistic face

Difference between Levels of Expressiveness

As shown in Table 9, there is no main effect of level of expressiveness in any variables. However, there are significant interactions between VOTE and perceived intelligence ($F(8.91, 2)$, $p<.01$) and comfortableness ($F(4.47, 2)$, $p<.05$). The AGREE group rated the Honest face as most intelligent, while the DISAGREE group rated the Stoic face as the most intelligent and the Honest face as the least. The AGREE group rated the Stoic face as least comfortable, while the DISAGREE rated the same face as most comfortable.

Table 9: The mean value of impressions for the Honest, Deceiving, and Stoic face in the game

Variables	Honest (n=20)	Deceiving (n=20)	Stoic (n=20)	F(value, df)
INT (Intelligence)	3.84	3.37	3.89	F(1.98, 2)
Vote Agree (n = 10)	4.20	3.10	3.30	F(8.91, 2)*
Vote Disagree (n = 10)	3.44	3.66	4.55	
LIKE (Likability)	4.37	4.37	4.68	F(1.28, 2)
ENG (Engagingness)	3.95	4.21	4.21	F(.56, 2)
COM (Comfortableness)	3.95	3.74	3.95	F(.14, 2)
Vote Agree (n = 10)	3.90	4.10	3.60	F(4.47, 2)**
Vote Disagree (n = 10)	4.00	3.33	4.33	
BEL (Correspondence)	4.79	4.79	4.63	F(6.44, 2)
ATT (Required attention)	4.32	4.47	4.37	F(1.46, 2)
DIS (distraction)	3.26	3.21	3.37	F(5.26, 2)
USE (usefulness)	4.21	4.42	4.11	F(1.81, 2)

** p<.01, ** p<.05, *** p<.10, no mark = not significant*

Observations, Interviews and Open Comments

Interviews and observations suggest that subjects tried to read the agent's poker playing strategy from its face. They attributed different personalities and characteristics to each face and thought each agent had a different playing strategy.

Subjects wrote many comments at the end of the questionnaire. Some subjects used "attitude" or "personality" to describe the characteristics of the players in their comments. Comments show that the subjects first tried to figure out whether the faces gave them clues on the player's hands and strategies. Then they noticed that the players' facial expressions gave information about their hands. This suggests that subjects used the faces to interpret the agents' behavior.

DISCUSSION

Subjects rated a poker playing agent with and without a face as equally intelligent. This suggests that attaching a face did not add any perceived intelligence to the poker player. However, having a face is considered more likable, engaging, and comfortable to play against regardless of subjects' opinion about personification. As described earlier, Takeuchi's [14] and Walker's [15] studies show a similar result, namely that having a face is engaging. It is encouraging that the face did not create a negative effect on the subjects' opinions about likability, even to those who have a negative opinion about personification, contrary to what Walker's study found [15]. This difference may be caused by the different nature of the tasks and faces used in these studies. Walker used a woman's synthetic talking face in a questionnaire survey, while this study used a caricature female face in a poker game environment.

Subjects' responses to the Human and the Dog's face, and Smiley, the Caricature, the Realistic face in a pilot experiment in which they evaluated the faces in isolation showed reasonable expectations from the appearance of each face. King's study [4] shows a similar result in perceived intelligence, namely that human forms are perceived to be more intelligent than other forms. King also suggests that fully articulated human forms are rated to have higher intelligence and agency than either caricatures or Chernoff faces (corresponding to the Smiley face in this experiment). However, subjects rated the same faces as equally intelligent (or unintelligent) in the poker game. For example, though people attributed different levels of perceived intelligence, likability, and engagingness to the Human face and the Dog's face, or Smiley, the Caricature, and the Realistic face when they evaluate the faces based on their appearance, there were no differences when they rated the same faces in the poker game. This means that they did not rate a player's intelligence based on its appearance. Their impressions are based on the player's actual competence in playing poker.

However, subjects' opinions about personification affect their attitude toward each of the faces. For example, in the HUMANITY experiment, the AGREE group are more "strict" about personification, since they prefer the Dog's face to the Human face, felt more comfortable with the Dog's face, and considered the Dog's face more appropriate for representing the actual competence of the player in playing poker. On the other hand, the DISAGREE group are more "flexible" about personification. The subject's gender works in the same way as the subjects' opinion about personification. Males

are more "strict", while female are more "flexible" about personification. However, due to the small number of female subjects, further study is needed to conclude that users' gender truly affects impressions of agent personification.

It is interesting that people's opinion about personification also affects the way they feel about expressiveness. The AGREE group rated expressive faces (Honest and Deceiving) to be more comfortable, and the Honest face to be most intelligent. The DISAGREE group rated the stoic face as most intelligent and comfortable. This means the DISAGREE group is more strict about simulating a real poker situation, where players are not expected to show honest expressions.

CONCLUSIONS AND FUTURE DIRECTIONS

The first finding is that having a face is considered more likable, engaging, and comfortable to see in a poker game environment. People did not feel distracted by the presence of a face or facial expressions. Moreover, people tried to interpret the faces and facial expressions, which makes the users pay attention to the face and engage in the task. Walker [15] describes the advantages and disadvantages of personification as follows: "This can lead to improved performance if the task is not very complex or to degraded performance if the task is complex." It is clear that faces are useful for entertainment purposes, since engagement is essential for games and people don't care about taking more effort for entertainment. Hence it may also be useful for applications which require an engaged user for success, such as education and training as evaluated in [1]. Andre [1] reports that the subjects rated learning tasks presented by an animated agent as less difficult than the presentations viewed without an animated agent.

The second finding is that people's impressions of a face are different when they see a face in isolation versus when they interact with a face within a task. People evaluate a face not based on appearance but on competence or performance. Most psychological HCI studies use static facial images separately from applications. One thing we learned from this study is that we have to evaluate effects of personification within a context--i.e. the context of software agent applications. As Don [3] says, the key issue for personification is to understand the nature of the task and the way of communication with the task. Of course not all software agents require personified interfaces. The goal of HCI work should be to understand when a personified interface is appropriate.

The third finding is that there is a dichotomy between user groups which have opposite opinions about personification. Differences in facial features such as character's gender, humanity, expressiveness cause opposite evaluations by these two subject groups. Some of the experiments performed indicate that there is also a potential difference in evaluation of a human face and a non-human face between the subjects' gender. Subjects' computer expertise and age range should also be taken into consideration. We need to consider the target users when designing a personified interface. The future personified interface should be flexible so that it can provide options to choose a preferred face or no face at all for each user.

REFERENCES

1. Andre, E. et al. WebPersona: a Lifelike Presentation agent for the World-Wide Web, Knowledge Based Systems, 2(1), 1998, 25-35
2. Bell, G. et al. Lifelike Computer Characters: The Persona project at Microsoft Research. Software Agents, edited by Jeff Bradshaw. Cambridge, MA: MIT Press, 1996.
3. Don, A. Anthropomorphism: From Eliza to Terminator 2, in Proceedings of CHI `92 (Monterey CA, May 1992), ACM Press, 67-70.
4. King, J. et al. The Representation of Agents: Anthropomorphism, Agency, and Intelligence, in Proceedings of CHI `96 (Vancouver, Canada, April 1996), ACM Press, 289-290.
5. Koda, T. Java Poker Game Experiment. Available as http://poker.www.media.mit.edu:81/epoker
6. Lashkari, Y. et al. Collaborative Interface Agents, in Proceedings of the National Conference on Artificial Intelligence (1994), Cambridge, MA: MIT Press
7. Laurel, B. Interface Agents: Metaphors with Character, in The Art of Human-computer Interface Design, edited by Brenda Laurel. Reading, MA: Addison-Wesley, 1990. 355-365.
8. Maes, P. Agents That Reduce Work and Information Overload, in Communications of the ACM. 37, 7 (July 1994), 31-40, 146.
9. Nass, C. et al. Computers as Social Actors, in Proceedings of CHI `94 (Boston MA, April 1994), ACM Press, 72-77.
10. Negroponte, N. Being Digital. The Alfred A. Knorpf, Inc. 1995. 101-102.
11. Oren, T. et al. Guides: Characterizing the Interface, in The Art of Human-computer Interface Design, edited by Brenda Laurel. Reading, MA: Addison-Wesley, 1990. 367-381.
12. Ortony, A. et al. The Cognitive Structure of Emotions. Cambridge, MA: Cambridge University Press. 1988.
13. Sheth, B. A Learning Approach to Personal Information Filtering. Master's Thesis. MIT Media Lab, 1994.
14. Takeuchi, A. et al. Situated Facial Displays: Towards Social Interaction, in Proceedings of CHI `94 (Boston MA, April 1994), ACM Press, 450-454.
15. Walker, J. et al. Using a Human Face in an Interface, in Proceedings of CHI `94 (Boston MA, April 1994), ACM Press, 85-91