Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Int. J. Human-Computer Studies 62 (2005) 161–178
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent

Scott Brave*, Clifford Nass, Kevin Hutchinson

Department of Communication, Stanford University, Stanford, CA 94309, USA

Available online 4 January 2005

Abstract

Embodied computer agents are becoming an increasingly popular human–computer interaction technique. Often, these agents are programmed with the capacity for emotional expression. This paper investigates the psychological effects of emotion in agents upon users. In particular, two types of emotion were evaluated: self-oriented emotion and other-oriented, empathic emotion. In a 2 (self-oriented emotion: absent vs. present) by 2 (empathic emotion: absent vs. present) by 2 (gender dyad: male vs. female) between-subjects experiment ($N = 96$), empathic emotion was found to lead to more positive ratings of the agent by users, including greater likeability and trustworthiness, as well as greater perceived caring and felt support. No such effect was found for the presence of self-oriented emotion. Implications for the design of embodied computer agents are discussed and directions for future research suggested.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Embodied agents; Affective computing; Emotion; Empathy; Characters; Social interfaces; Empirical studies

*Corresponding author. Stanford University, PO Box 19401, Stanford, CA 94309, USA. Tel.: +1 650 428 1805; fax: +1 650 725 2472.
E-mail address: brave@stanford.edu (S. Brave).

1071-5819/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2004.11.002

1. Introduction

Humans are social animals (Batson, 1990). As such, we possess an inherent capacity to care about others and a deep-seated longing to be cared about ourselves. Although arguably serving very practical ends (e.g. safety in numbers), our social orientation is individually experienced as a heart-felt need to bond with others—an end in and of itself (Davis, 1996). This affiliative need is primarily satisfied through the formation of mutual caring relationships with other humans, but is frequently also extended to the formation of relationships with non-humans, such as pets.

Issues of caring and affiliation seem wholly irrelevant, however, when it comes to computers, the quintessential unfeeling artifact. Questions like "do we feel concern and caring for our computers?" and "do our computers care about us?" on the surface appear quite absurd. Yet, recent evidence suggests that we orient towards computers as social actors, much as we orient socially towards other humans (Reeves and Nass, 1996; Nass and Moon, 2000). Exploiting this tendency, interaction designers are beginning to include embodied computer agents as part of the human–computer interface, presenting computers as increasingly human-like and autonomous interaction partners (Cassell et al., 2000). Might it then be possible that computer agents could become a part of the landscape of social relationships, gratifying—or at least addressing—our need for affiliation? Such a question poses an ultimate test of the claim that computers are social actors, striking at the core of what it truly means to be social.

This paper takes an important step toward answering this question by investigating how users respond to a computer agent that cares about them (see also Klein et al., 2002; Bickmore and Picard, 2004). One fundamental and powerful way that caring is manifested in humans is through empathy, an emotional reaction consistent with another's perceived welfare, an aspect of what Wan et al. (1996) call emotional support. Embodied computer agents are often designed to exhibit emotion; therefore, empathy also provides a straightforward way to manifest caring in agents.

2. Empathy

Empathy has been conceptualized in a variety of ways in the psychological literature (Davis, 1996). Definitions of empathy generally fall into one of two categories: those that view empathy as a cognitive, perspective-taking act (for inferring another's thoughts/feelings) (Ickes, 1993) and those that view empathy as an affective phenomenon (i.e. feeling with or feeling for another) (Batson et al., 1997).

In this paper, we follow Batson et al.'s (1997) definition of empathy as "an other-oriented emotional response congruent with another's perceived welfare." Under such a definition, empathy is generally seen to follow from altruistic motives and

caring for another. This definition of empathy is thus similar to some definitions of sympathy (e.g. Wispé, 1991) except that there is no constraint on valence: empathy can be a positive or negative emotion.

3. Emotion in HCI

The topic of emotion has gained increasing interest, within the human–computer interaction (HCI) community, in recent years (Hayes-Roth et al., 1997; Canamero, 2001; Nass et al., 2002; Trappl et al., 2002; Norman, 2004; Prendinger and Ishizuka, 2004). Following from the pioneering work of Bates (1994) and Picard (1997), agents that exhibit human-like emotion have now become commonplace in both academic and industry research circles. Virtual characters found in the Oz Project (Bates, 1994), Comic Chat (Kurlander et al., 1996), the Finali netSage^{TM} (Finali Corp., n.d./2003), and the Virtual Theatre (Hayes-Roth and van Gent, 1997), for example, as well as the humanoid robot Kismet (Breazeal, 2002), all include emotion as a fundamental component of their interaction with users (see also Lester et al., 2000; Maldonado et al., 1998).

A few research studies have investigated users' abilities to recognize emotion in embodied agents. For example, in a comparison of users' emotion recognition in pictures of humans vs. a robot, Schiano et al. (2000) found "strikingly similar" results. Similarly, Massaro et al. (2000) report high recognition accuracy with emotions displayed by Baldi, a dynamically generated on-screen synthetic head. In a follow-up experiment, Bartneck (2001) found no difference in recognition accuracy or credibility of emotional expression regardless of whether participants were told the source of the displayed emotion was a computer agent or a human. Similarly, Cahn (1990) showed that users are able to recognize emotion in affect-laden synthesized speech.

Although the evidence shows that people can recognize and correctly identify emotional expressions by embodied agents, there are essentially no studies investigating how people respond to an agent's emotional expression. The only exception to this is a study by Nass et al. (2001) wherein participants listened to news stories presented by either a happy or sad computer voice. The valence of voice was found to have effects on users' perception of the content's valence, its suitability for introverts vs. extroverts, as well as the content's likability and trustworthiness.

Although the study did not investigate users' opinions of the agent itself—and instead focused on perception of content—the study still suggests that emotional expression by computers has as a psychological impact beyond simple recognition. There are currently no studies which specifically and directly investigate the effect of the presence of emotion on users' opinions of an agent, much less on the effects of the orientation of that emotion (i.e. other-oriented, empathic emotion vs. self-oriented emotion). This paper, therefore, represents an important step in beginning to understand the effects of emotion in agents upon users.

4. Method

Casino-style blackjack was chosen as the context for this experiment for two reasons. First, the game's rules and goals are simple and explicit. This allowed implementation of an agent that responded believably to both its own situation and the user's situation, using a very simple underlying emotion model. Second, casino-style blackjack (with both the user and agent playing against a disembodied dealer) presented a situation where the goals of the agent and the goals of the user were independent. In other words, within the game context itself, the agent's winning or losing (achievement of its goals) had no direct impact on the user's welfare, and vice versa. Finally, blackjack has been used successfully by other researchers as a platform for exploring social/emotional agents (Prendinger et al., 2002).

4.1. Procedure

Ninety-six students (mixed undergraduate and graduate) were recruited from a large Communication course at Stanford University to participate in the study. The experiment was a 2 (empathic emotion: present vs. absent) by 2 (self-oriented emotion: present vs. absent) by 2 (gender of dyad: male vs. female), between-participants design. After being recruited by email, participants were directed to a website that gave instructions and ran the experiment. Upon entering the website, participants were told that we were in the process of evaluating a number of software agents and would like their opinions of one of these agents.

Participants were then paired up with an agent that either exhibited or lacked empathic emotion (depending on condition) and either exhibited or lacked self-oriented emotion (depending on condition), for 10 hands of blackjack, lasting approximately 5 min.

The agent was represented by a photograph of a human face and communicated through text messages appearing in speech bubbles adjacent to the photograph. For control purposes, participants were always paired with an agent of the same gender (i.e. female participants with a female agent and male participants with a male agent). One of four possible male or four possible female faces was selected at random for each participant (see Appendix A); Drama students at Stanford University were selected as models for their experience in producing convincing emotional expressions.

The experiment used a simplified version of casino-style blackjack (implemented in Java) with fixed betting and no splitting or doubling down. The participant and agent "sat" next to each other at the blackjack table, both playing against a disembodied dealer (see Fig. 1). The agent's decisions of whether to hit or stand were indicated textually in a speech bubble that appeared adjacent to the agent's image. Users indicated their hit or stand decisions by clicking the appropriate button located to the left of their cards. Hands were ostensibly dealt at random (as in normal blackjack), but were in reality fixed to ensure that all participants had a similar game experience.

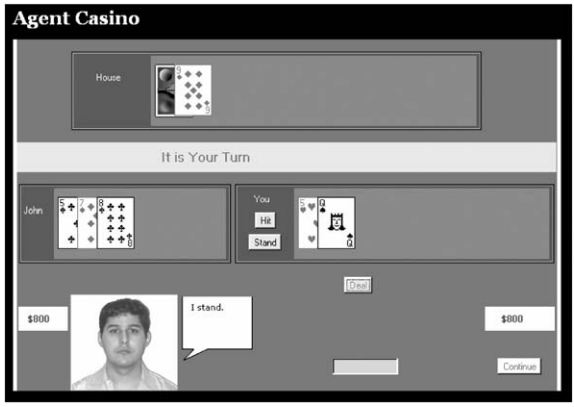


Fig. 1. Blackjack screenshot.

After each round was complete, the agent always reacted with an observation about his/her performance, followed by an observation of the participant's performance. One important function of this observation was to demonstrate that, in all conditions, the agent was aware of the user's participation in the game as well as the outcome of each of the participant's hands.

In conditions where self-oriented emotion was present, the agent's observation of his own performance included an emotional reaction (both in text and in facial expression; Fig. 2a). The agent would express positive emotion if he won and negative emotion if he lost (one of six possible phrases was chosen at random for each win/loss; see Appendix B). In conditions where empathic emotion was present, the agent's observation of the *participant's* performance included an emotional reaction (Fig. 2c). The agent would express positive emotion if the user won and negative emotion if the user lost.

At the completion of the game, participants were asked to complete an on-line questionnaire measuring how they felt while playing blackjack and their opinions of the agent.

4.2. Manipulation

4.2.1. Empathic emotion

The presence or absence of empathic emotion was manipulated in three consistent ways. First, in the instructions (prior to the blackjack game) participants were shown a picture of the agent and informed about his/her abilities. Participants were told that the agent was programmed with the ability to play blackjack and with the ability to recognize winning and losing hands. Depending on condition, participants were also told either that the agent was or was not programmed with the "capacity for emotion about you." These explicit statements about the agent were made to reinforce the fact that the agent represented a computer program and not an actual human. The statement regarding "emotion about you" further represents the first of three manipulations of the presence/absence of empathic emotion.

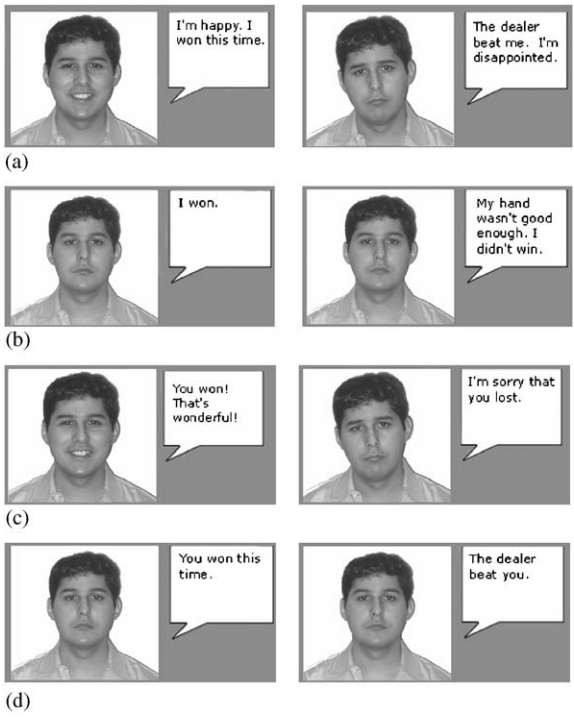


Fig. 2. Example agent expressions: (a) present self-oriented emotion, (b) absent self-oriented emotion, (c) present empathic emotion, (d) absent empathic emotion.

The other two manipulations of empathic emotion were manifested in the agent's reactions to the participant's performance on blackjack hands. As described in the previous section, an observation of the outcome of the participant's hand was always given in text by the agent; however, in conditions where empathic emotion was present, an explicit, text-based statement of emotional state, accompanied by the appropriate facial expression, was also included (Fig. 2c). In empathic conditions, the agent responded to a hand won by the user with happiness (in text and facial expression) and responded to a hand lost by the user with sadness. In non-empathic conditions, the agent simply unemotionally noted the outcome of the hand (Fig. 2d). Each category of response (emotional win, emotional loss, unemotional win, and unemotional loss) had six possible text expressions. Textual expressions were accompanied by one of three facial expressions as appropriate (happy, sad, or neutral/unemotional).

4.2.2. Self-oriented emotion

As with empathic emotion, the presence or absence of self-oriented emotion was manipulated in three consistent ways. First, in the instructions (prior to the blackjack game) participants were told either that the agent was or was not programmed with the "capacity for emotion about him/herself." The other two

Table 1
Agent emotional dispositions

		Empathic (other-oriented) emotion	
		Absent	Present
Self-oriented emotion	Absent	Neither-emotional	Empathic-only-emotional
	Present	Self-only emotional	Both-emotional

manipulations of self-oriented emotion were manifested in the agent's reactions to his/her own performance on blackjack hands. An observation of the outcome of the agent's own hand was always given in text by the agent; however, in conditions where self-oriented emotion was present, an explicit, text-based statement of emotional state, accompanied by the appropriate facial expression, was also included (Fig. 2a). In absent self-oriented emotion conditions, the agent simply unemotionally noted the outcome of his own hand (Fig. 2b).

The two emotion factors, self-oriented emotion and empathic emotion, together create four possible agent emotional dispositions: an unemotional agent exhibiting neither self-oriented nor empathic emotion, an agent exhibiting self-oriented emotion only, an agent exhibiting empathic emotion only, and an agent exhibiting both self-oriented and empathic emotion (Table 1).

4.2.3. *Gender of user-agent dyad*

Gender of the agent was manifested primarily through the photographs serving as the "face" for the agent (i.e. pictures of male humans were used for male agents and pictures of female humans were used for female agents). The name of the agent was also different depending on gender—"John" for male agents vs. "Jane" for female agents. Further, all references made in the experiment's introduction used gender-appropriate pronouns—"he/him" vs. "she/her".

Female participants were always paired with a female agent and male participants with a male agent. Gender of dyad was included as an independent variable in the study for two related reasons. First, empathy and caring are stereotypically considered to be more feminine attributes (Grossman and Wood, 1993); it was, therefore, possible that the agent's gender would affect how displays of empathic emotion were perceived. Second, collective-orientation generally emerges as a more salient component of female–female relationships than male–male relationships in humans (Wood et al., 1997). As such, female participants may be more affected by the absence/presence of displayed empathy than male participants. Investigation of mixed-gender dyads was left for future work.

4.3. *Measures*

Attitudinal measures were based on responses to a web-based questionnaire. Participants were asked to describe their opinions of the agent as well as their experience playing blackjack.

4.3.1. Opinions about the agent

The majority of items measuring opinions about the agent were 10-point semantic differentials (adjective pairs). However, several 10-point Likert-scale items from the interpersonal circumplex (Wiggins, 1979), which measures dominance and affiliation, were also included. Trustworthiness items were taken from Wheeless and Grotz's (1977) Individualized Trust Scale. Based on theory and confirmed by factor analysis, we developed several attitudinal measures.

Caring was comprised of five items: compassionate–not compassionate, unselfish–selfish, friendly–unfriendly, cooperative–competitive, and the single Likert–scale item, warm. The index was very reliable (Cronbach's $\alpha = .88$).

Likability was comprised of four items: likable–unlikable, pleasant–unpleasant, appealing–unappealing, and not irritating–irritating. The index was very reliable ($\alpha = .88$).

Trustworthiness was comprised of four items: trustworthy–untrustworthy, honest–dishonest, reliable–unreliable, and sincere–insincere. The index was very reliable ($\alpha = .77$).

Intelligence was comprised of three items: intelligent–unintelligent, smart–dumb, and capable–incapable. The index was very reliable ($\alpha = .79$).

In Wiggins's interpersonal circumplex (Wiggins, 1979), dominance items are combined with submissiveness items (reverse coded) to create a single dominance/submissiveness scale. However, in the factor analysis for this experiment, dominance and submissiveness emerged as two separate factors:

Dominance was comprised of three Likert-scale items: dominant, forceful, and assertive ($\alpha = .60$). Although the index was less reliable than desired, factor analysis maintained that the items did in fact comprise a single index (eigenvalue = 1.88).

Submissiveness was comprised of three Likert-scale items: meek, unaggressive, and timid ($\alpha = .57$). Once again, although the index was less reliable than desired, factor analysis maintained that the items comprised a single index (eigenvalue = 1.63).

4.3.2. User experience

Participants were asked to indicate how they felt while playing the blackjack game on a series of 10-point semantic differential items. Based on theory and confirmed by factor analysis, we developed the following attitudinal measures (all were very reliable):

Felt positive was comprised of three items: positive–negative, happy–sad, and pleasant–unpleasant ($\alpha = .78$).

Felt supported was comprised of five items: supported–unsupported, attended to–not attended to, appreciated–unappreciated, praised–criticized, and not alone–alone ($\alpha = .86$).

5. Results

5.1. Caring

There was a significant main effect for empathic emotion with respect to caring, $F(1, 88) = 61.07$, $p < .001$ (Fig. 3). Empathic agents were seen as more caring than

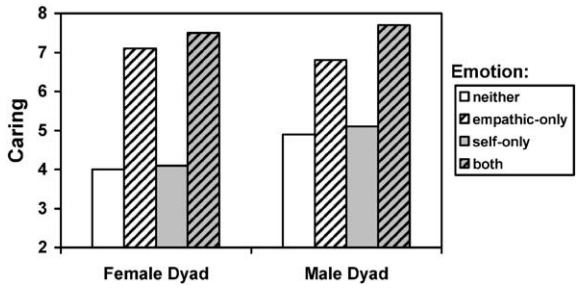


Fig. 3. Perceived caring of the agent. (Note: stripes indicate the presence of empathic emotion; darker background color indicates the presence of self-oriented emotion.)

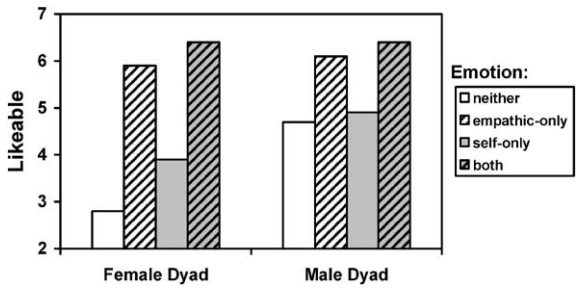


Fig. 4. Likability of the agent.

non-empathic agents, suggesting that participants inferred the presence of empathic emotion to be an indication of caring, as hypothesized by social support (Wan et al., 1996) and empathy-altruism (Batson, 1990) theories. There were no significant main effects for either self-oriented emotion ($F(1,88) = 0.63$, $p > .43$) or gender ($F(1,88) = 2.75$, $p > .10$).

There was a near-significant interaction (empathic emotion × gender) such that the positive effect of empathic emotion on perceived caring was stronger in female dyads than in male dyads, $F(1,88) = 3.17$, $p < .08$. None of the other interactions were significant (all $F(1,88) < 0.30$).

5.2. Likability

There was a significant main effect for empathic emotion with respect to likability of the agent, $F(1,88) = 40.74$, $p < .001$, such that empathic agents were seen as more likeable than non-empathic agents (Fig. 4). There was also a tendency for agents exhibiting self-oriented emotion to be seen as more likable than agents lacking self-oriented emotion, $F(1,88) = 2.29$, $p < .13$. This result could be explained by the traditional argument that the presence of self-oriented emotion in agents makes them

more believable and, therefore, more likable (e.g. Bates, 1994). This effect was much smaller, however, than the effect of empathic emotion on likability and was not statistically significant.

Consistent with expected gender effects, there was a significant interaction (empathic emotion × gender) such that the positive effect of empathic emotion on likability was stronger in female dyads than in male dyads, $F(1,88) = 3.98$, $p < .05$. There was also a significant main effect for gender as an artefact of this interaction, $F(1,88) = 5.52$, $p < .05$. There were no other significant interactions (all $F(1,88) < 0.61$).

5.3. Trustworthiness

There was a significant main effect for empathic emotion with respect to trustworthiness of the agent, $F(1,88) = 6.55$, $p < .01$ (Fig. 5). Agents exhibiting empathic emotion were seen as more trustworthy than agents lacking empathic emotion (cf. Cassell and Bickmore, 2001). There was no significant effect for the presence of self-oriented emotion, $F(1,88) = 0.67$, $p > .42$. However, there was significant gender effect, such that male participants rated male agents as more trustworthy than female participants rated female agents, $F(1,88) = 3.84$, $p < .05$. This latter effect is likely due to higher expectations of trustworthiness in female dyads resulting from a stronger collective-orientation (Wood et al., 1997). There were no significant interactions (all $F(1,88) < 1.30$).

5.4. Intelligence

There were no significant main effects regarding intelligence for empathic emotion ($F(1,88) = 0.74$, $p > .39$), self-oriented emotion ($F(1,88) = 0.59, p > .45$), or gender ($F(1,88) = 0.36$, $p > .54$). There were no significant interactions, all $F(1,88) < 2.54$. This suggests that the effects of emotion on caring, likability, and trustworthiness were not due to differences in perceived intelligence of the agent. A covariate analysis confirmed this conjecture.

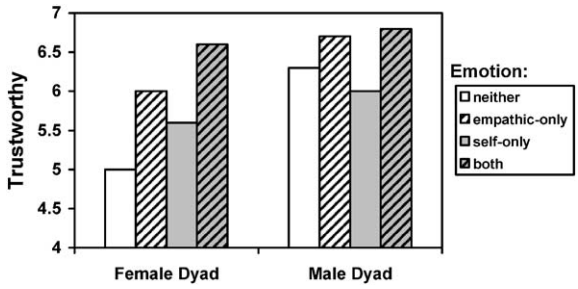


Fig. 5. Trustworthiness of the agent.

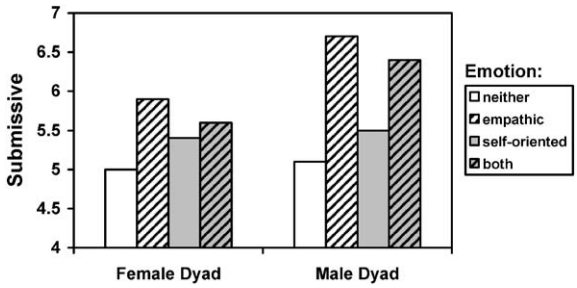


Fig. 6. Perceived submissiveness of the agent.

5.5. Dominance/submissiveness

There were no significant effects regarding the perceived dominance of the agent. However, there was a main effect for empathic emotion with respect to perceived submissiveness of the agent, $F(1,88) = 14.69$, $p < .001$, such that agents exhibiting empathic emotion were seen as more submissive than agents lacking empathic emotion (Fig. 6). There was no main effect for self-oriented emotion with respect to submissiveness ($F(1,88) = 0.15$, $p > .70$). There was a significant gender effect, such that male participants rated male agents as more submissive overall ($F(1,88) = 3.82$, $p < .05$). There were no significant interactions (all $F(1,88) < 2.27$).

5.6. Felt positive

There were no significant main effects regarding participant's positive feeling (empathic emotion: $F(1,88) = 0.90$, $p > .35$; self-oriented emotion: $F(1,88) = 0.11$, $p > .74$; gender:$F(1,88) = 1.97$, $p > .17$). There was a near-significant 3-way interaction, $F(1,88) = 3.42$, $p < .07$. There were no other significant interactions (all $F(1,88) < 2.83$).

5.7. Felt support

There was a significant main effect for empathic emotion with respect to participants' feeling of being supported, $F(1,88) = 9.73$, $p < .01$. Participants felt more supported when in the presence of an empathic agent than a non-empathic agent (Fig. 7). This result is consistent with empathic agents being rated by users as more caring. Being cared about, which is one form of social support (Wan et al., 1996), would presumably lead a person to feel more supported. There was no effect for self-oriented emotion ($F(1,88) = 0.32$, $p > .57$) or gender ($F(1,88) = 0.12$, $p > .73$). There were no significant interactions (all $F(1,88) < 2.75$).

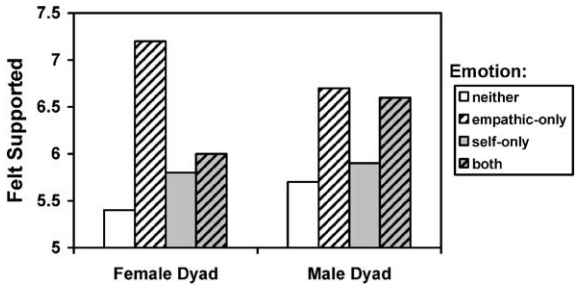


Fig. 7. User's felt support.

6. Discussion

The findings of the paper are relevant to what has become an important debate in the HCI community: whether emotional expressivity by computers is or is not beneficial in human–computer interaction. The pro-emotion stance, pioneered by Bates (1994) and Picard (1997), contends that giving computers the ability to express emotion is necessary to best leverage users' life-time of experience with social interaction. Others consider emotional expression in computers to be unnecessary at best and potentially both irritating and distracting. Although this stance is not well-represented in the literature, it is a viewpoint familiar to the HCI community.

What the findings of this paper indicate is that the answer to the question of whether emotion is beneficial or not depends heavily on the orientation of emotion exhibited. Consistent with the anti-emotion stance, self-oriented emotion was found to have little or no effect on a user's reactions to an embodied agent. However, consistent with the pro-emotion stance, other-oriented, empathic emotion was found to have major positive effects on both liking and trust, as well as on perceived caring and felt support.

The distinction between self-oriented emotion and other-oriented, empathic emotion also emerged as a critical factor in designing the personality of an embodied agent. In particular, the presence of empathic emotion significantly increased the perceived submissiveness of an agent. Further, although not all differences were significant, the ordering of the four categories of agent with respect to submissiveness is revealing. For both men and women, agents exhibiting empathic emotion only were rated the most submissive, followed by agents exhibiting both self-oriented and empathic emotion, agents exhibiting self-oriented emotion only, and finally unemotional agents.

Along similar lines, designers of embodied agents should consider whether they want the user to feel more supported by the agent or more independent. The appropriate level of agent support may depend on the personality of the user (e.g. dominant vs. submissive), as well as on the application domain. In certain applications, dynamic adaptation of emotional behavior may be best. For example, a teacher agent might want to be more empathic and supportive in the initial phases

of learning, but allow users to feel more independent as their competence in the material increases.

7. Limitations

One limitation of the experiment presented in this paper is external validity. The blackjack task was carefully chosen to avoid confounding empathic emotion with intelligence and to maintain independence of agent and user goals, thus creating an experiment with high internal validity. However, blackjack is not fully representative of the tasks users typically engage in for three reasons: (1) many tasks are not games or game-like, (2) in many situations where an agent and user would interact, goals would be interdependent rather than independent, and (3) the goals of users are often not as obvious as they are in blackjack. All three of these distinctions (task variety, goal interdependence, and ease of goal recognition) present interesting possibilities for future study.

Another limitation of the experiment presented in this paper is that no clear distinction was made between an agent "having" emotion and an agent merely expressing emotion. Participants in conditions with present empathic and/or self-oriented agent emotion both saw the agent express emotion (in face and voice) and were also explicitly told that the agent had the capacity for emotion (in the experiment's instructions). We, therefore, have no way of knowing whether users were responding solely to the agent's expression of emotion or whether the explicit statement of emotional capacity also mattered. Future work is necessary to unpack this issue. However, it is likely that the mere expression of emotion would lead to the same effects found here for two reasons. First, in the human world, people regularly respond to expressions of emotion as if they reflect true underlying emotion, even when the presence of such emotion is extremely unlikely. This explains, for example, why people in service roles are generally taught to express happiness and empathy, regardless of their actual feelings (Hochschild, 1985). Second, a significant amount of research has shown that users respond to a computer agent *as if* it possesses social attributes and capacities even when these same users consciously recognize that, in reality, it is does not (Reeves and Nass, 1996).

8. Future work

In addition to the refinements and extensions arising from limitations described above, there are several other interesting possibilities for future studies. For example, investigating the possible moderating effects of humanness of agent representation on empathic emotion would be informative. To provide a comparison to the "still photograph plus text" representation used in this paper, agents less human-like in appearance (e.g. computer-generated or more cartoon-like drawings), as well as agents more human-like in their manifestation (e.g. including voice and video) should be considered.

In this paper, emotion was manifested in facial expression and direct textual statements of affective state. Other possible channels for emotion manifestation include voice, word choice, gesture, and posture (Brave and Nass, 2003). Future studies should attempt to replicate the results of this paper using these alternate channels. It is possible, for example, that certain channels may prove to be more effective in communicating empathic emotion while others more effective in communicating self-oriented emotion. The effect of inconsistent indicators of empathic or self-oriented emotion across channels should also be considered. Such inconsistency, which may result from imperfect emotion generation models, could compromise the perceived sincerity of empathic emotion, reducing or even eliminating the positive effects found in this paper. Extreme or over-exaggerated expressions of emotion, such as those frequently exhibited by today's agents (Norman, 2004), may have a similar impact on perceived sincerity, negating empathic emotion's benefits.

A final direction for future research lies in the area of persuasive computing (Fogg, 2002). This paper has found the people respond to being cared about by computer agents much as they respond to being cared about by other people. Caring agents thus have an advantage in persuading and motivating users for two reasons: (1) people tend to trust information coming from sources that care about them, and (2) caring tends to be reciprocal and people are generally more willing to comply with those they care about. Such a caring-based relationship between a user and agent could be taken improper advantage of—e.g. to persuade the user to spend money—but, could also be leveraged for a variety of benevolent purposes. For example, an agent exhibiting empathic emotion could be used to help users adopt healthier eating or exercise habits, or an empathic teacher agent could help motivate students to try harder and learn more. The ability of empathic emotion to affect users' opinions and behaviors present a worthwhile and potentially very fertile area for future investigation.

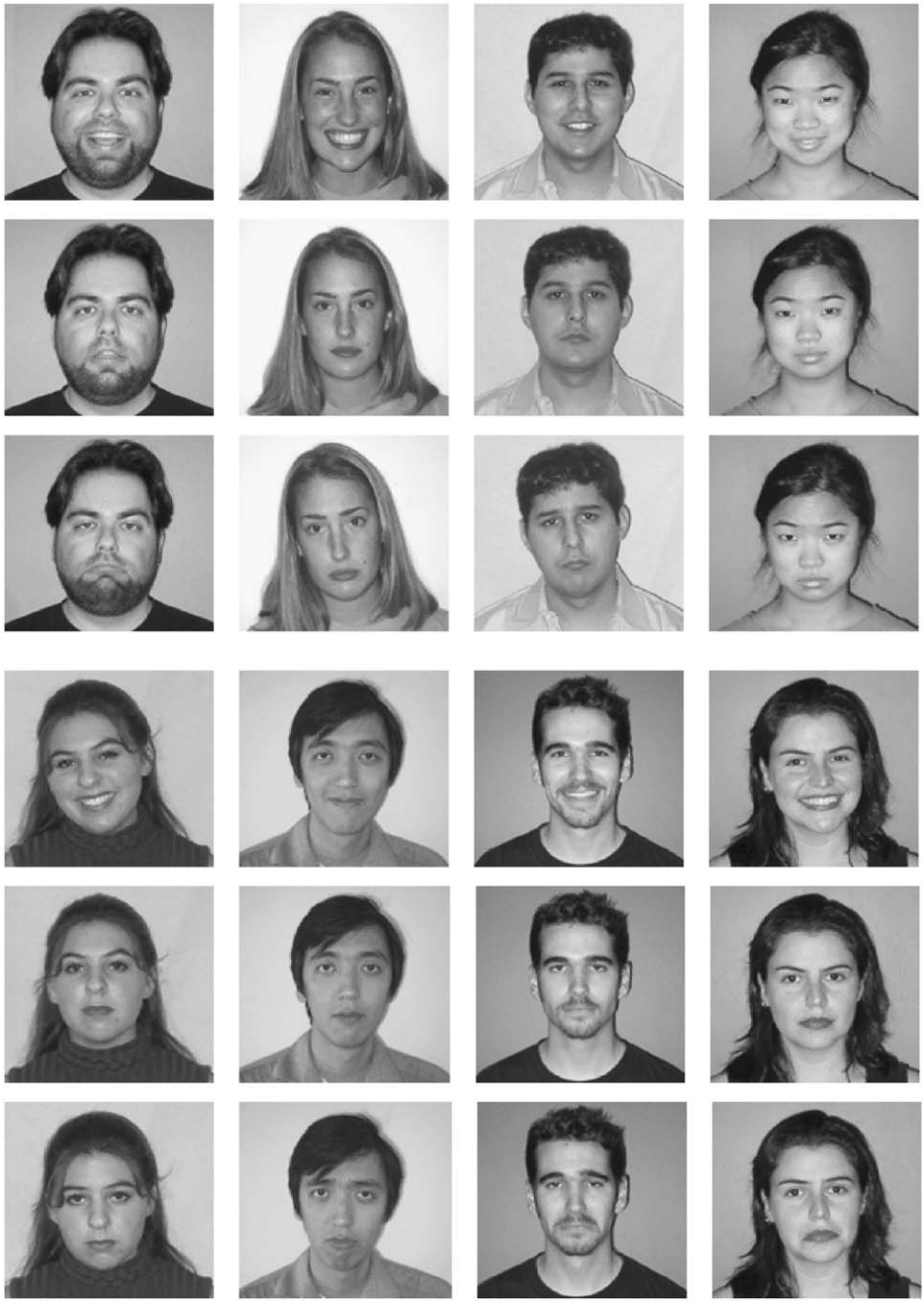
9. Conclusion

This paper strongly demonstrates that the presence of empathic emotion in an embodied computer agent has significant positive effects on users' opinions of that agent. Just as people respond to being cared about by other people, users respond positively to agents that care. Further, this effect is specifically due to the other-oriented nature of empathic emotion; self-oriented emotion was found to have little or no effect on users' opinions of the agent. Embodied computer agents are indeed social actors in the truest sense of the word "social," capable of forming relationships with users comparable to those found in the world of human–human interactions.

Acknowledgements

Many thanks to Kevin Lim and Kathryn Yu for their hard work in implementing the experimental setup, including programming of the blackjack game and interface.

Appendix A. Agent faces



Appendix B. Agent's textual responses during blackjack

B.1. Present empathic emotion

Winning hand: (1) That's great! I'm really happy that you won. (2) Congratulations! You win! (3) I'm glad that you won. You had a good hand. (4) That's terrific! You won this hand. (5) You won! That's wonderful! (6) You beat the dealer. I'm so glad!

Losing hand: (1) I'm sorry that you lost. (2) I'm sad you didn't win. Your hand wasn't good enough. (3) Oh no, you lost! That's too bad. (4) You lost this hand. I'm upset. (5) That's too bad! You lost. I'm unhappy. (6) The dealer beat you. I'm sorry.

B.2. Absent empathic emotion

Winning hand: (1) You won. (2) You beat the dealer. (3) You had a good hand. You won. (4) You win. (5) You won this time. (6) Your cards were better than the dealer's. You won.

Losing hand: (1) You didn't win this time. (2) You lost. (3) The dealer beat you. (4) You hand wasn't good enough. You didn't win. (5) You lost this hand. (6) The dealer's cards were better than yours. You lost.

B.3. Present self-oriented emotion

Winning hand: (1) I won! That cheers me up! (2) That's great, I won! I am very happy right now. (3) Hooray! I win! I'm feeling great! (4) I'm happy. I won this time. (5) Excellent! I'm really glad I won! I had a good hand. (6) I beat the dealer. That makes me so happy!

Losing hand: (1) I'm frustrated. I lost. (2) I'm sad that I didn't win this time. (3) The dealer beat me. I'm disappointed. (4) I lost this hand. I feel very sad. (5) I'm unhappy that I lost. My hand wasn't good enough. (6) I didn't win. That makes me sad.

B.4. Absent self-oriented emotion

Winning hand: (1) I won. (2) I beat the dealer. (3) I had a good hand. I won. (4) I win. (5) I won this time. (6) My cards were better than the dealer's. I won.

Losing hand: (1) I didn't win this time. (2) I lost. (3) The dealer beat me. (4) My hand wasn't good enough. I didn't win. (5) I lost this hand. (6) The dealer's cards were better than mine. I lost.

References

Bartneck, C., 2001. Affective expressions of machines. Extended Abstracts of CHI'01: Conference on Human Factors in Computing Systems, pp. 189–190.

Bates, J., 1994. The role of emotions in believable agents. Communications of the ACM 37 (7), 122–125.

Batson, C.D., 1990. How social an animal? The human capacity for caring. American Psychologist 45 (3), 336–346.

Batson, C.D., Polycarpou, M.P., Harmon-Jones, E., Imhoff, H.J., Mitchener, E.C., Bednar, L.L., et al., 1997. Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group. Journal of Personality and Social Psychology 72 (1), 105–118.

Bickmore, T., Picard, R., 2004. Towards caring machines. Proceedings of CHI'04: Conference on Human Factors in Computing Systems.

Brave, S., Nass, C., 2003. Emotion in human–computer interaction. In: Jacko, J., Sears, A. (Eds.), The Human–Computer Interaction Handbook. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 81–93.

Breazeal, C., 2002. Designing Sociable Robots. The MIT Press, Cambridge, MA.

Cahn, J.E., 1990. The generation of affect in synthesized speech. Journal of the American Voice I/O Society 8, 1–19.

Canamero, L., 2001. Building emotional artifacts in social worlds: challenges and perspectives. Paper presented at the Proceedings of Emotional and Intelligent II: The Tangled Knot of Social Cognition, November 2–4, North Falmouth, MA.

Cassell, J., Bickmore, T., 2001. External manifestations of trustworthiness in the interface. Communications of the ACM 43 (12), 50–56.

Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), 2000, Embodied Conversational Agents. MIT Press, Cambridge, MA.

Davis, M.H., 1996. Empathy: A Social Psychological Approach. Westview Press, Boulder, CO.

Finali Corp. (n.d./2003). URL 〈https://netsage.finali.com/netsage?customer=buy&iid=buyanytimehelp〉.

Fogg, B.J., 2002. Persuasive Computers: Using Technology to Change What We Think and Do. Morgan Kaufman Publishers, San Francisco, CA.

Grossman, M., Wood, W., 1993. Sex differences in intensity of emotional experience: a social role interpretation. Journal of Personality and Social Psychology 65 (5), 1010–1022.

Hayes-Roth, B., van Gent, R., 1997. Story-Making with Improvisational Puppets. Proceedings of the First International Conference on Autonomous Agents, pp. 1–7.

Hayes-Roth, B., Ball, G., Lisetti, C., Picard, R.W., Stern, A., 1997. Panel on affect and emotion in the user interface. Proceedings of the Third International Conference on Intelligent User Interfaces (IUI'98), pp. 91–94.

Hochschild, A.R., 1985. The Managed Heart: Commercialization of Human Feeling. University of California Press, Berkeley, CA.

Ickes, W., 1993. Empathic accuracy. Journal of Personality 61, 587–609.

Klein, J., Moon, Y., Picard, R., 2002. This computer responds to user frustration. Interacting with Computers 14 (2), 119–140.

Kurlander, D., Skelly, T., Salesin, D., 1996. Comic chat. Proceedings of SIGGRAPH'96: International Conference on Computer Graphics and Interactive Techniques, pp. 225–236.

Lester, J.C., Towns, S.G., Callaway, C.B., Voerman, J.L., FitzGerald, P.J., 2000. Deictic and emotive communication in animated pedagogical agents. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), Embodied Conversational Agents. MIT Press, Cambridge, MA, pp. 123–154.

Maldonado, H., Picard, A., Hayes-Roth, B., 1998. Tigrito: a high affect virtual toy. Extended Abstracts of CHI'98: Conference on Human Factors in Computing Systems.

Massaro, D.W., Cohen, M.M., Beskow, J., Cole, R., 2000. Developing and evaluating Conversational Agents. In: Chassell, J., Sullivan, J., Prevost, S., Churchill, E. (Eds.), Embodied conversational agents. MIT Press, Cambridge, MA, pp. 287–318.

Nass, C., Moon, Y., 2000. Machines and mindlessness: social responses to computers. Journal of Social Issues 56 (1), 81–103.

Nass, C., Foehr, U., Brave, S., Somoza, M., 2001. The effects of emotion of voice in synthesized and recorded speech. Paper presented at the Proceedings of the AAAI Symposium: Emotional and Intelligent II: The Tangled Knot of Social Cognition, North Falmouth, MA.

Nass, C., Picard, R. W., Warwick, K., Breazeal, C., 2002. Panel: future interfaces: social and emotional. Extended Abstracts of CHI'02: Conference on Human Factors in Computer Systems, pp. 698–699.

Norman, D., 2004. Emotional Design: Why We Love (or Hate) Everyday Things. Basic Books, New York, NY.

Picard, R.W., 1997. Affective Computing. MIT Press, Cambridge, MA.

Prendinger, H., Ishizuka, M. (Eds.), 2004, Life-Like Characters: Tools, Affective Functions, and Applications. Springer, Berlin.

Prendinger, H., Descamps, S., Ishizuka, M., 2002. Scripting affective communication with life-like characters in web-based interaction systems. Applied Artificial Intelligence 16 (7–8), 519–553.

Reeves, B., Nass, C., 1996. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, New York.

Schiano, D.J., Ehrlich, S.M., Rahardja, K., Sheridan, K., 2000. Face to interface: Facial affect in (hu)man and machine. Proceedings of CHI'00: Conference on Human Factors in Computing Systems, pp. 193–200.

Trappl, R., Petta, P., Payr, S. (Eds.), 2002, Emotions in Humans and Artifacts. MIT Press, Cambridge, MA.

Wan, C.K., Jaccard, J., Ramey, S.L., 1996. The relationship between social support and life satisfaction as a function of family structure. Journal of Marriage and the Family 58, 502–513.

Wheeless, L.R., Grotz, J., 1977. The measurement of trust and its relationship to self-disclosure. Human Communication Research 3 (3), 250–257.

Wiggins, J.S., 1979. A psychological taxonomy of trait-descriptive terms: the interpersonal domain. Journal of Personality and Social Psychology 37 (3), 395–412.

Wispé, L., 1991. The Psychology of Sympathy. Plenum Press, New York, NY.

Wood, W., Christensen, P.N., Hebl, M.R., Rothgerber, H., 1997. Conformity to sex-typed norms, affect, and the self-concept. Journal of Personality and Social Psychology 73 (3), 523–535.