

Affective social robots[☆]

Rachel Kirby^{a,*}, Jodi Forlizzi^b, Reid Simmons^a

^a Carnegie Mellon University, Robotics Institute, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

^b Carnegie Mellon University, Human–Computer Interaction Institute and Design Department, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:

Received 25 July 2007

Received in revised form

4 September 2009

Accepted 21 September 2009

Available online 30 September 2009

Keywords:

Human–robot interaction

Social robots

Emotions

Moods

Affective modeling

ABSTRACT

For human–robot interaction to proceed in a smooth, natural manner, robots must adhere to human social norms. One such human convention is the use of expressive moods and emotions as an integral part of social interaction. Such expressions are used to convey messages such as “I’m happy to see you” or “I want to be comforted,” and people’s long-term relationships depend heavily on shared emotional experiences. Thus, we have developed an affective model for social robots. This generative model attempts to create natural, human-like affect and includes distinctions between immediate emotional responses, the overall mood of the robot, and long-term attitudes toward each visitor to the robot, with a focus on developing long-term human–robot relationships. This paper presents the general affect model as well as particular details of our implementation of the model on one robot, the Roboceptionist. In addition, we present findings from two studies that demonstrate the model’s potential.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Social robots, such as those that operate in healthcare institutions and in museums, need to communicate with people in ways that are natural and easily understood, even by non-roboticists. We believe that one way to improve these interactions is to have robots display changing moods and emotions, just as humans do. This paper describes a generative model of affect that attempts to strongly mimic how people emote in order to produce as natural-seeming a system as possible. The model is designed particularly for robots that interact with people over long periods of time. As such, our focus is on modeling the long-term aspects of, and interactions between, emotions, moods, and attitudes. We have implemented our affective model on the Roboceptionist, a robot that interacts with people on a daily basis [1]. In addition, we have run several experiments to demonstrate the model’s use in social situations, which show that people *do* recognize emotional expressions on the robot’s face (Fig. 1) and that such expressions can significantly influence how people interact with the robot.

1.1. Human interaction

Affect, such as mood and emotion, plays a major role in human interaction. Quite often, emotional reactions are caused

by social interactions, influenced by societal and cultural norms, or used to communicate desires to other people [2]. Emotions carry conversational content, allowing conversational partners to form common ground and communicate more effectively [3]. For instance, an expression of sadness—facial, vocal, or behavioral—may indicate a desire to be comforted. Furthermore, what mood a person is in has a strong impact on how that person interacts with others [4]; for example, people who are interacting may “catch” each other’s moods and emotions, unconsciously matching their own emotional states to their conversational partners’ [5]. Frijda argues that the primary reason for social interaction is, in fact, to experience emotions, which serve to form a “sense of coherence with others” [6]. Suppression of emotions can be highly detrimental to relationship forming and is disruptive to conversations [7].

A well-studied effect of human–computer interaction is that people tend to treat computer agents in the same way that they treat other people, forming social relationships with them [8]. We believe that this tendency to form social relationships with computers will also apply to robots, perhaps even more so. If that is the case, then people will respond to a robot’s emotions as though the robot were human, and will expect the robot’s emotional responses to be consistent across multiple interactions.

1.2. Human–robot social interaction

In recent years, the robotics community has seen a gradual increase in social robots, that is, robots that exist primarily to interact with people. Museum tour-guide robots [9] and robots that interact with the elderly [10] demonstrate not only the

[☆] This work was funded in part by an NSF Graduate Research Fellowship to the first author and by NSF grants #IIS-0329014 and #IIS-0121426. Portions of this work have appeared previously in Gockley et al. (2006) [34,37].

* Corresponding author. Fax: +1 412 268 5571.

E-mail addresses: rachelg@cs.cmu.edu (R. Kirby), forlizzi@cs.cmu.edu (J. Forlizzi), simmons@cs.cmu.edu (R. Simmons).

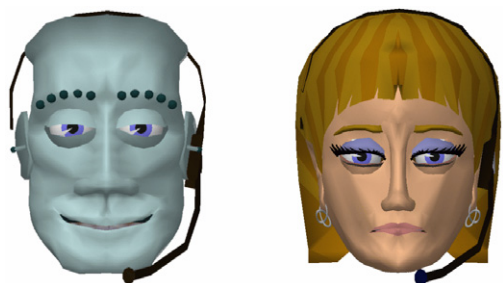


Fig. 1. Two of the graphical faces used on the Roboceptionist. On the left is Tank displaying happiness; on the right is Valerie displaying sadness.

benefits of having robots interact with people, but also the need for the interactions to be smooth and natural. Many of these robots have incorporated at least some rudimentary emotional behaviors, but such behaviors are usually ad hoc and not generalizable to other robots. Robots with infant-like abilities of interaction, such as Kismet [11], have been used to demonstrate the ability of people to interpret and react appropriately to a robot's displays of emotions. Experiments with the robot Vikia [12] demonstrated the effectiveness of an emotionally expressive graphical face for encouraging interactions with a robot. All of these, however, have focused on short-term interactions with the robot. We believe that a richer model of affect is necessary for forming long-term human–robot relationships.

The exploration of affect in social robots currently lags behind similar research in software agents. For example, the Affective Reasoner [13] is an implementation of a virtual world populated by software agents, wherein the agents can detect and react to each other's emotions; however, the agents do not interact directly with humans. In contrast, Embodied Conversational Agents (ECAs) are designed explicitly to mimic human–human interaction [14]. Many ECAs are capable of expressing emotions (see [15] for an overview). One such system, called FearNot!, has been shown to invoke empathy in its users in response to the agents' emotional displays [16]. However, the emotional models used in these types of systems are typically ad hoc, incomplete, or poorly documented, and thus difficult to extend to other systems. Many systems rely on a computational model of emotions developed by Ortony, Clore, and Collins [17], but again, this model does not account for affective behavior beyond very short-term emotional responses. In contrast, the system described by El Jed and colleagues [18] includes longer-term moods and personality in addition to emotions, but we believe that our model provides more human-like behavior for long-term interactions, as we will discuss in the following sections.

The most developed affective model specifically for robots that we are aware of at this time is the TAME architecture [19], which considers the four categories of personality traits, attitudes, moods, and emotions. Our model follows a similar breakdown of categories (namely, attitudes, moods, and emotions). However, the TAME model does not account for the interplay between different affective phenomena, nor does it specify how the robot's affect may be influenced by its interactions with people. In contrast, we have tried to provide a complete model of affect with a strong psychological backing, designed specifically for long-term human–robot interaction.

2. Affect

As background, this section provides a brief description of affect and how it relates to human interaction. **Affect** is a general term relating to emotions, moods, and other such states with varying degrees of positivity or negativity—that is, states with *valence*. While a great deal of psychological research has

focused on affect, only a few researchers distinguish between terms such as “emotion” and “mood,” and very few agree on the meanings of these terms. Following the categorization suggested by Scherer [20], we consider the following phenomena to be of importance in developing an affective model: emotions, moods, and attitudes.

2.1. Emotion

An **emotion**, or “emotional response,” is an immediate affective response to the evaluation of some event (or other stimuli) as being of major significance. While psychologists may debate whether emotions have distinct facial responses (see e.g. [21,22]), when considered solely for interaction purposes emotions are meaningless unless they result in some outward change in the robot, including facial, vocal, or behavioral modifications.

In the psychological and cognitive science literature, there are two primary views on the representation of emotions: categorical and continuous. Ekman [23] and others argue for a set of “basic” emotions that are innate and universal across cultures. All other emotional categories are then built up from combinations of these basic emotions. Others, such as Russell [24], argue that all emotions lie in a continuous two-dimensional space, where the dimensions are typically taken to be *valence* (how positive or negative the emotion is) and *arousal* (the energy or excitation level associated with the emotion). Both representations have been used in robotic applications; for example, Sage [9] uses a categorical model, whereas Kismet [11] uses a continuous model. In both cases, people identified the robots' emotions and reacted appropriately; the underlying representation does not seem to have a major effect on people's understandings of a robot's emotional displays. In our work, we employ a categorical model of emotions, as discussed in Section 3.1.

2.2. Mood

Moods are more “diffuse” affective states [20] that typically do not have a single antecedent. They are typically of lower intensity than emotions and have fairly low variance over the course of a single day. Moods may be caused by a variety of things, including changes in physiological state (such as lack of sleep or illness) [25], rapidly occurring emotional responses [25], or complex cognition regarding emotional life events [26].

Several studies have indicated that positive moods tend to reduce negative emotions (in frequency and intensity), and negative moods tend to reduce positive emotions [25,27]. Additionally, evidence from psychology indicates that daily moods tend to have plateaus, during which repeated emotional events will not cause a large shift in mood [28]. In particular, mood due to life stresses is fairly stable and is not strongly influenced by most emotional reactions. However, many similar emotions in a row will eventually alter the mood significantly.

Interacting with other people can have a strong impact on one's mood, but the exact effects of such social interactions are still not well understood. Some evidence indicates an asymmetrical crossover model, wherein positive social exchanges will increase a positive mood but have little effect on a negative mood, while negative social exchanges will erode any mood [29].

2.3. Attitude

An **attitude** is an amalgamation of emotions experienced with a particular person (or thing), reflecting one's relationship with that person over time. Three primary characteristics of attitudes are that they are learned over time, they predispose one to certain behaviors, and they imply evaluations on an affective scale [30].



Fig. 2. A visitor interacting with Valerie, the Roboceptionist. The robot's virtual head is displayed on a monitor that rotates to face visitors. The robot's body—an RWI B21r—is "dressed" in a suit.

Changes in attitudes may result from emotions inspired by the focus of the attitude, co-experienced emotional events, or the frequency and duration of the relationship. A person's attitude toward a conversational partner can influence many aspects of the conversation; for example, people are less likely to express emotions in the presence of strangers [31]. Attitudes are a key part of long-term relationships.

Emotions, moods, and attitudes all interact. In particular, both emotions and attitudes help to shape the overall mood, and, at the same time, the mood effects how strongly emotional reactions occur. Additionally, changes in mood during social interactions result in changes of attitude. The exact mathematical interactions used in our affective model are described below.

3. Implementation

As discussed, several models of human affect have been postulated, and which model is "correct" remains undetermined. However, we require only that our robot's behavior is human-like; the underlying model need not precisely match human cognition. That is, we wish to design a generative model, rather than an explanatory one. Accordingly, we have taken the approach of selecting a simple, straightforward model of human affect to serve as the basis of our computational model.

Our current platform for social robotics research is the Roboceptionist [1]. The mechanical base of the robot is an RWI B21r with an LCD "head" that can rotate to face visitors. The robot is housed in a booth near a high-traffic entrance to a computer science building at Carnegie Mellon University, and is generally available 8 hours a day, 5 days a week. A keyboard and small monitor on the booth's desk allow for human input with visual feedback; speech recognition is not used due to high environmental noise. The LCD head displays a highly expressive, graphical face, which can display a wide range of easily recognizable emotional expressions. In a collaboration with the Drama Department, the robot has gone through two character iterations, each with a unique background and a complex, evolving storyline. Visitors to the booth can ask the robot questions about its life in order to hear the continuing story, according to which the first Roboceptionist, Valerie (shown in Fig. 2), left her job to become

a singer and was replaced with Tank, a paranoid ex-CIA agent. These stories are intended to provide interest and to encourage the formation of long-term relationships with the robot.

3.1. Emotions

We chose to implement a categorical model of emotions, primarily due to the relative ease of implementing several distinct emotional expressions versus defining a mapping between continuous multidimensional space and the robot's expressions. In particular, we implemented a subset of the "basic" emotions suggested by Ekman [23]. These emotions are: joy (happiness), sadness, disgust (frustration), and anger, as shown in Fig. 3. While this list is in no way comprehensive, additional emotions (such as fear or surprise) can be added easily, as desired. Each emotion has an associated intensity level, represented as a real number ranging from 0 (non-existent) to 1 (highest intensity), as well as a valence rating (positive or negative). For each emotion, we defined a series of expressions of differing intensities, which can be displayed on the robot's graphical face. Specific intensities are generated by a linear interpolation of the muscle positions between the two nearest defined intensity expressions (see example in Fig. 4). The emotional expressions used were based on Delsarte's code of facial expressions as implemented for the robot Vikia [12].

Since this model is intended for social robots, emotions are caused primarily by interactions that the robot has with people. For example, a new person interacting with the robot may cause happiness, and insults typed to the robot may generate sadness or anger. We have implemented a mechanism within the robot's language model in order to trigger specific emotions directly. That is, certain statements by visitors carry specific emotional content, such as compliments causing happiness or insults resulting in sadness. Currently, the emotional intensities are hand-coded, but a deeper model could be implemented using the cognitive model proposed by Ortony, Clore, and Collins [17], in which emotions occur as the agent evaluates events, objects, or other agents. Bartneck [32] provides more information on how this model could be used.

Emotions are displayed immediately after an event, and last the duration of the robot's verbal response. In this way, emotions are short-lived, but are displayed long enough to be recognized. Emotions do not occur concurrently in our model, thus avoiding the question of how different base emotions might interact or "blend." This greatly simplifies the display of emotions. Since emotions are short-lived, multiple emotions can be displayed sequentially.

3.2. Moods

Our robot's moods are primarily caused by its personal history and "life" events. That is, because the robot has an ongoing life story, it can feel positive (or negative) about past (or future) events. Values for the moods are assigned to the storyline by the dramatic writers, as they see fit. The mood generated from these events is considered as the robot's "baseline" mood for each day. The robot's overall mood is influenced by the emotions it experiences throughout the day, as explained below. In different robots, moods may also be influenced by other circumstances, such as internal power levels or the ability to complete assigned tasks. Our robot's moods are indicated by posture, particularly the tilt of the head (e.g. a downcast face indicates a negative mood). A more sophisticated indication of mood might also include vocal and behavioral modifications.

After a life event has occurred, that event's contribution to the robot's mood fades over time. Events that cause more intense moods take longer to fade than lesser events. Similarly, anticipated

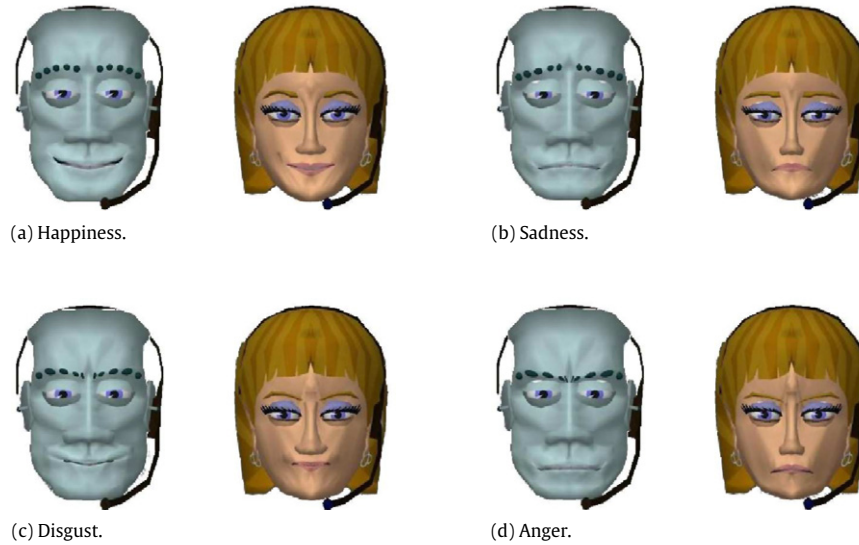


Fig. 3. Four emotional expressions, shown on both Tank and Valerie's faces. Both faces have a similar range of expressions. Each emotion is displayed at its highest intensity.

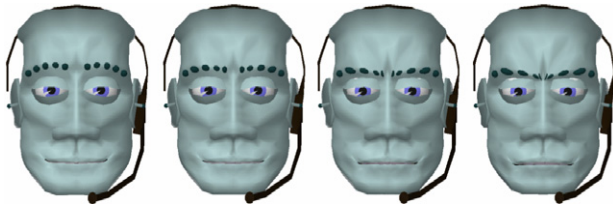


Fig. 4. Interpolation of emotional expressions between neutral (far left) and very angry (far right).

events will increase in salience as the date of the event approaches, causing an increase in the mood contribution in the days preceding the event. Specifically, if the event will occur in d days and will cause a mood of strength m on the day it occurs, then that event's contribution c to the mood is computed according to the sigmoid curve given below:

$$c = m - \frac{m}{1 + e^{-|d|+|10m|}} \quad (1)$$

Events that have already occurred follow the same curve. Some examples can be seen in Fig. 5. Multiple events' contributions to the mood are additive.

In our model, the robot's mood has one primary plateau [28], at the baseline mood as computed for each day. As discussed in the next sections, various events throughout the day may create an offset from the baseline, causing the mood to change dynamically. We constrain the mood to change smoothly and to approach the extremes of the facial expressions asymptotically, so that the model will never attempt to generate an expression outside the normal bounds of the facial muscles. To model this effect, the robot's displayed mood follows a generalized logistic (growth) curve, as given in the following equations:

$$m_d = \begin{cases} \frac{1 + m_b}{1 + e^{-B(m_o + M(1+m_b))}} - 1 & \text{if } m_o < 0 \\ m_b & \text{if } m_o = 0 \\ \frac{1 - m_b}{1 + e^{-B(m_o - M(1-m_b))}} + m_b & \text{if } m_o > 0 \end{cases} \quad (2)$$

where m_d is the displayed mood, m_o is the offset from the baseline mood, and m_b is the baseline mood. These equations follow a logistic curve with a growth rate of B ; this influences how quickly

the mood changes between the plateau and each extreme. Our current implementation uses $B = 15$, but this value can be modified, if necessary, for specific robots. The time of maximum growth of each half-curve (that is, the inflection points) occurs at $M(1 \pm m_b)$, where in our implementation $M = 0.35$. This point is scaled according to the baseline to keep the rate of change of the mood similar across different baselines. These equations may be better understood by viewing some examples for different baseline moods, shown in Fig. 6. While the baseline mood and offset may be any real numbers, the curves given in Eq. (2) constrain the displayed mood to a range of -1 (extremely negative) to $+1$ (extremely positive), with 0 representing a neutral mood.

Variables such as B and M can be thought of as representing the robot's "personality traits." Varying their values will alter how long the robot remains in the baseline plateau (M) and how quickly its mood changes outside of the plateau (B).

3.3. Interaction between mood and emotion

To model human behavior, the robot's mood level should modulate the intensities of its emotions [25,27]. In our implementation, we model this effect by scaling the strength of the emotional response linearly, according to the current mood. Mood-congruent emotions (that is, emotions that have the same valence as the mood) are increased in intensity, while incongruous emotions are diminished. Specifically, if an emotion of nominal strength s ($0 \leq s \leq 1$) and valence v ($v = \pm 1$) occurs during a mood of m , then the emotion's strength is scaled according to Eq. (3):

$$s' = s \left(1 + \frac{1}{4}vm \right) \quad (3)$$

The exact scaling may be moderated by a particular robot's personality traits [33].

In addition to the mood influencing emotional responses, emotions in turn influence the mood. In our implementation, the robot's mood dynamically changes due to its interactions with people. For example, the robot may be in a negative mood due to some event in the storyline, but can be "cheered up" through repeated happy interactions during the day.

As discussed above, positive social exchanges increase positive moods while negative exchanges erode any mood [29]. If the robot experiences a negative emotion during any mood or a positive

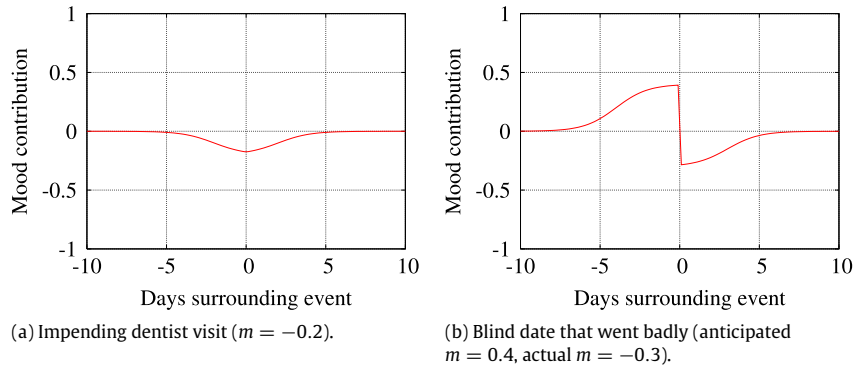


Fig. 5. Examples of different events and how they contribute to the overall mood over several days. Each event occurs at day 0.

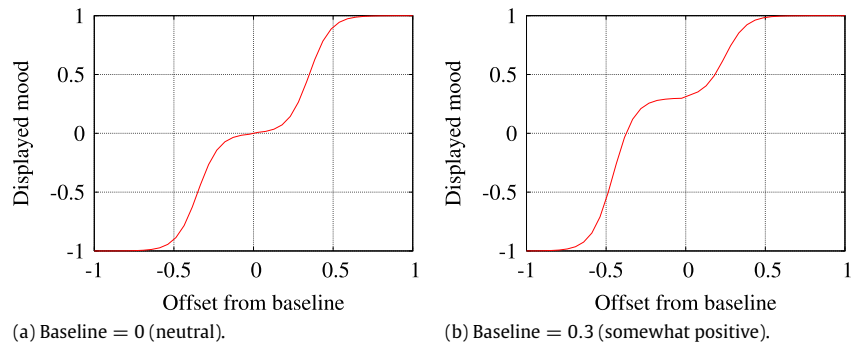


Fig. 6. How the displayed mood changes based on the underlying mood level, expressed as an offset from the “baseline” mood. Offsets result from changing emotions throughout the day.

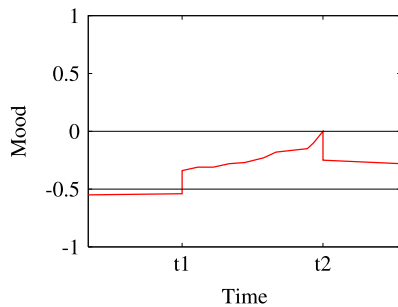


Fig. 7. An example of how attitude and interactions influence mood. Here, the baseline mood is -0.5 , and a previous interaction has lowered the mood even further. A well-liked person ($A_m = 0.4$) approaches to begin interaction at time t_1 , which immediately improves the robot’s mood. The person interacts from t_1 to t_2 , largely trying to cheer up the robot, improving its mood considerably. When the person leaves, the mood drops once again, but remains at a higher level than before. After the interaction, the mood begins to decay toward the baseline.

emotion during a positive mood (i.e. $m_b + m_o > 0$), the mood offset (m_o) changes by one fifth the strength of the emotion—a change that is small enough to be unnoticeable after only a small number of exchanges, but may cause a significant change in the robot’s expression (see Eq. (2)). Positive exchanges during a negative mood have little effect, increasing the mood offset by only half what other exchanges cause (that is, one tenth the strength of the emotion).

While the overall mood changes dynamically due to interactions with people during the day, the “baseline” mood is designed to dominate the robot’s mood. That is, after a person has stopped interacting with the robot, any effect that the person had on the robot’s mood will begin to decay. Currently, we implement this decay by decreasing the mood offset (m_o) by small increments at regular intervals (currently, 10% every 30 s). Thus, if there are no further interactions over time, the mood will eventually return to the baseline level.

3.4. Mood and attitude

An attitude is essentially a long-term mood associated with some person (or thing). Each person who visits the robot may cause various emotional responses in the robot, thus changing the robot’s mood, and those changes influence the robot’s “opinion” of that person. How well the robot “knows” the person may influence how the robot’s attitude changes toward that person. For the robot to maintain believability, its attitude toward someone should remain consistent across multiple interactions with that person.

In our model, attitudes consist of a mood level (A_m) and a familiarity rating (A_f). Familiarity is computed by combining how many hours a person has ever spent interacting with the robot and how many days since the person last visited the robot. Because we could find no psychological evidence for a computational model of familiarity, we have chosen a simple linear combination of these two values, as given in Eq. (4). A person has the highest possible familiarity rating (1) if she has interacted with the robot for at least 10 hours (including all past interactions with the robot), and the last time she interacted with the robot was that same day. Visitors are assigned familiarity ratings as follows:

$$A_f = \frac{1}{2} \left(1 + \frac{1}{10} \min(\text{hours}, 10) - \frac{1}{30} \min(\text{days}, 30) \right). \quad (4)$$

A stranger is assumed to have interacted infinitely many days ago, and so his familiarity rating is 0.

When a person begins an interaction with the robot, the robot’s attitude toward that person has an immediate effect on its mood. Specifically, the mood offset changes to the average of the current offset (m_o) and the attitude’s mood level (A_m). Strangers to the robot are given an attitude level of 0, which will pull the robot’s mood closer to the baseline plateau—making the robot less likely to display emotions to strangers. When a person leaves the robot, the robot’s mood offset becomes the average of its mood offset

User	Good afternoon
Tank	Welcome to Newell Simon. [happy 0.1]
User	I am a new student here
Tank	Do you enjoy your job?
User	Yes
Tank	That's good. I like my new job here.
User	What's your job
Tank	I can do all sorts of tasks. I can give directions. I can help you find a room or a faculty or staff member in the School of Computer Science. I know the current weather for all the major cities of the world.
User	Could you please tell me where is Dr. [name]?
Tank	[name] is in [room number] Newell Simon Hall. Would you like directions?
User	Yes, please
Tank	You are already on the third floor of Newell Simon Hall. Go across the atrium... [further directions]
User	Do you know where I can find a map
Tank	Sorry. I don't know who or what that is
User	ok
Tank	[nods]
User	thanks
Tank	You're welcome. [happy 0.4]
User	Have a nice day
Tank	Glad to have helped. Have a pleasant day.
User	bye
Tank	Thanks for visiting Newell Simon. Come back and see me again
User	sure :)
Tank	I didn't understand that. It might help if you rephrase that. Or, ask me something simple. [disgusted 0.5]

Fig. 8. An interaction between the robot and a novice user.

immediately before the person arrived and its mood offset when the person leaves. Furthermore, the robot's attitude toward that person changes by a fraction of the robot's mood change during the interaction Eq. (5). Higher levels of familiarity result in a more stable attitude across interactions. An example of how a person can alter the robot's mood can be seen in Fig. 7.

$$A'_m = A_m + (\Delta m_o \text{ during interaction})(1 - A_f). \quad (5)$$

4. Interactions

While it is difficult to determine what might constitute a "typical" interaction with the Roboceptionist, the following examples may provide some insight into how interactions may progress. Both examples are from actual visitors to the Tank the Roboceptionist.

4.1. New visitor

The text of this interaction is shown in Fig. 8. In this interaction, which occurred in August of 2009, the visitor identified himself as new to the University. This visitor triggered three emotional responses from the robot: two instances of happiness at being helpful, and one instance of disgust/frustration at not understanding the user's response. The robot's mood began as slightly negative before the interaction ($m_d = -0.017$), peaked at $m_d = 0.017$,

returned to neutral ($m_d = 0$) at the end of the interaction, and finally returned to a somewhat negative value after the visitor left ($m_d = -0.017$).

4.2. Emotional

This interaction, shown in Fig. 9, occurred in January of 2009. In this interaction the visitor invoked multiple emotions in the robot. This visitor was recognized by the robot, and had the associated attitude of $A_f = 0.50$ and $A_m = 0$. Before the interaction, the robot's mood was neutral $m_d = 0$. During the interaction, the robot's mood drops to a low of -0.10 after the extremely angry response (according to Tank's backstory, its real name is "Marion LaFleur," which it hates). After the person leaves, the robot is left in a slightly negative mood ($m_d = -0.02$), and the robot's attitude toward the person has dropped to ($A_m = -0.09$). Note that the robot's current interaction model does not maintain context of the interaction, so that each time the visitor typed "lol" (an abbreviation for "laughing out loud"), the robot reacts with amusement, improving its mood—even though a human would likely have experienced annoyance. A better dialog manager could improve the consistency of the robot's emotional responses.

5. Analysis

We conducted experiments to test components of our affect model. In particular, we verified that people can recognize the

User	idiot
Tank	How rude. I am not used to that type of abuse. [angry 0.5]
User	lol
Tank	I'm glad you are amused. [happy 0.3]
User	marion
Tank	I've told you people: DON'T CALL ME MARION! [angry 0.8]
User	lol
Tank	I'm glad you are amused. [happy 0.3]

Fig. 9. A short interaction with the robot that triggered multiple emotional responses.

robot's expressions of different emotions at various intensities, and we have studied the effect of the robot's apparent mood on how people interact with it.

5.1. Emotions

We ran a short web-based survey to verify whether people are able to “read” the robot's emotional expressions.

5.1.1. Procedure

In the survey, participants were asked to watch a series of video clips of a graphical face (either Valerie or Tank, see Fig. 1) displaying either a neutral expression or a single emotional expression (happiness, sadness, anger, or disgust) in various intensities (half-strength or full intensity). The expressions for different intensities of each emotion were computed as discussed in Section 3.1; the muscle motions to express each emotion were identical across robot faces. Expressions were shown in random order, and each participant viewed all 18 video clips. Each clip lasted about 1 s and showed the graphical face changing from a neutral expression to the emotional expression (or in the case of the neutral clips, just small eye and head movements). Participants could replay the videos as often as they wished. For each video, participants were asked to rate the robot's valence and arousal, on scales of 1–9. Note that participants were *not* asked to assign emotion words to the video clips; they were asked only for these valence and arousal ratings. Additionally, participants were asked some basic demographic information, including gender and whether they had ever interacted with the Roboceptionist.

5.1.2. Results

While over 200 people completed at least part of the survey, we report only on the 110 respondents who completed the entire survey. Roughly 58% of the participants were male, and nearly all (95%) were between the ages of 18–34. Of the participants, 61% stated that they had interacted with the Roboceptionist (either Tank or Valerie) on at least one occasion.

To analyze whether participants could distinguish between the various intensities of each emotion, we ran four separate two-way (robot \times emotion intensity) multivariate analyses of covariance (MANCOVAs), with valence and arousal as the dependent variables. Additionally, all analyses controlled for the participant's gender and previous interaction with the Roboceptionist. The “neutral” clip was used as intensity 0 for each emotion. Participants' ratings of the different emotions at different intensities can be seen in Fig. 10. Emotional intensities were compared by two orthogonal contrasts: neutral versus the low intensity expression, and the high intensity expression versus both neutral and low intensity. For all emotions, there was a significant main effect of intensity on participants' ratings of the robot's valence and arousal (F 's > 16.7 , p 's < 0.0005). While both contrasts were useful in differentiating

the conditions, the contrast of high intensity versus the lower intensities contributed more strongly to the discriminant function. Which robot was viewed also had a significant effect (F 's > 9.94 , p 's < 0.0005); in general, Tank was rated as higher in both valence and arousal than Valerie.

Similarly, we ran a two-way (robot \times emotion type) MANCOVA, again controlling for gender and prior interaction, to determine whether participants could differentiate between the different emotional categories. For this analysis, we considered only the responses to the high intensity emotional expressions, due to their greater discriminatory power as found in the previous analyses. We defined three orthogonal contrasts to explore the differences between the emotions: a valence contrast (happy versus the three negative emotions), a passivity contrast (sad versus angry and disgusted), and a final comparison between similar negative emotions (angry versus disgusted). We found that the differences between the emotions were significant ($F[6, 101] = 69.8$, $p < 0.0005$). Individual univariate tests indicated that valence best differentiated between happiness and the negative emotions and between anger and disgust, while arousal best differentiated between sadness and the other negative emotions (p 's < 0.0005). The robot had a significant effect on participants' responses ($F[2, 105] = 15.2$, $p < 0.0005$), following a similar trend as mentioned in the previous analysis, with Tank receiving higher valence and arousal ratings than Valerie.

5.1.3. Discussion

This study showed that people are able to detect differences between the robot's emotional expressions as well as differentiate between the intensities of each emotion. Furthermore, the valence and arousal ratings for each intensity followed roughly linear trends in the appropriate directions for each emotion (e.g. increasing happiness resulted in increasing valence and arousal, while increasing sadness resulted primarily in decreasing valence). We believe that this validates our model's use of interpolation between differing intensities of emotions.

It is interesting to note that, while both faces use the same underlying musculature structure and the same muscle position definitions for each emotion, participants rated the different faces as significantly different. It may be that the more mechanical face (Tank) is simply seen as more “pleasant” or appropriate than the more human-like face (Valerie), or that the gender of the robot's face causes different interpretations. Thus, for the model to be consistent across different robots, each robot's emotional expressions may need to be adjusted until they are seen as equivalent by people interacting with the robots. This could be done by repeating the study presented here in order to analyze which expressions of emotion are viewed as equivalent across different robots.

Furthermore, because this study was run on-line, participants viewed only the graphical face of the robot, but not the robot itself. Clearly, the robot's physical embodiment has a strong influence on people's perceptions. However, we do not expect that people's recognition of the emotional expressions will change significantly when seen on the robot rather than on a separate display, as the facial expressions will remain the same.

Finally, we note that this study served primarily to validate the linear interpolation of emotional expressions. While it is important that participants were able to identify the specific expressions, this aspect of the study does not easily generalize to other robots (particularly those with mechanical faces). Each implementation of a particular emotional expression must be validated. However, such validation need be done only for the most extreme intensity of each emotion, as lower intensities can be reliably interpolated.

5.2. Moods

While we have shown that people are able to recognize the robot's emotional expressions, our model focuses strongly on the

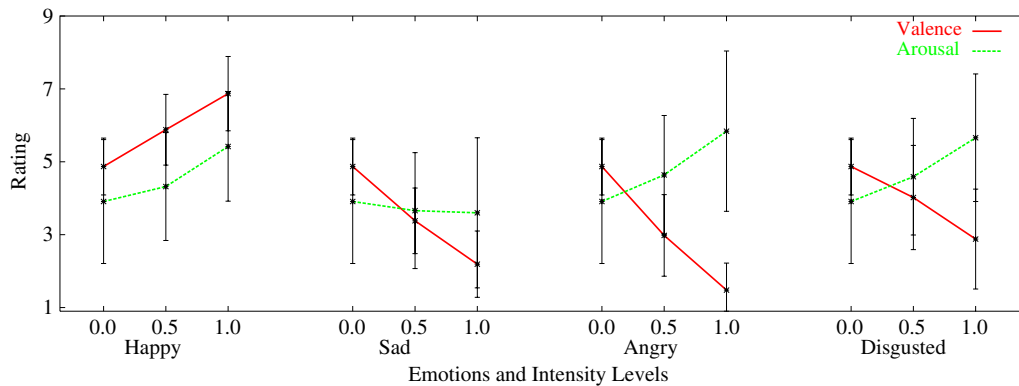


Fig. 10. Mean participant ratings for three intensities of each emotion, averaged across the robot face condition. Note that both valence and arousal follow linear trends with respect to intensity, and these trends differ for each emotion. Intensity level 0 for each emotion corresponds to the same video of “neutral,” and thus is identical for each emotion. Error bars indicate standard deviations. $N = 110$.



Fig. 11. The three expressions used to test the effect of the robot’s mood: neutral, positive, and negative (left to right). Differences include eyebrow and eyelid positions, mouth shape, and head angle.

effect of the robot’s long-term mood. Thus we have run a study to determine how the mood of the Roboceptionist (as Valerie) influences people’s interactions with it during a longer-term study.

5.2.1. Procedure

The study was run for a total of nine weeks, with the robot typically operating eight hours per day, five days per week. During each day, the robot displayed either her usual “neutral” mood or a positive or negative mood. We altered the moods across days and weeks both to match the robot’s life stories and to account for daily differences in interactions. The different moods were indicated by a change in facial posture, as shown in Fig. 11, as well as minor behaviors the robot performed during conversations while waiting for the visitor to type a statement. These behaviors included:

- In the neutral condition, Valerie either smiled and performed a single head nod or briefly glanced away from the visitor. These were the same behaviors that Valerie typically performed prior to this study.
- In the positive condition, Valerie either smiled and nodded as with neutral (but with a wider smile), or bounced her head from side to side in a seemingly happy, energetic motion.
- In the negative condition, Valerie either looked away from the visitor or appeared to sigh.

Within each condition, the different behaviors were selected randomly with equal likelihood during pauses in the interactions. We pre-tested the facial expressions and behaviors and found that people could accurately determine the intended moods. To avoid bias due to incongruent mood displays, the positive mood was run during weeks in which the storyline was also positive, and the negative mood during weeks in which the storyline was negative. No other changes were made to the interaction structure. As described elsewhere [1], people can interact with Valerie by typing on a keyboard, and Valerie responds using automatic speech generation. Visitors have the option of swiping an ID card in Valerie’s cardreader to identify themselves to the robot, but the

robot otherwise has no means of identifying interactors even if they had previously interacted with the robot.

During each day of the experiment, we used the laser rangefinder in the booth to detect and record the number of people who interacted with the robot. For each interactor, the robot logged the following: the time the interaction began, the number of seconds from the start of interaction until the visitor left the laser’s detection range (about two meters away from the keyboard), the number of lines of input typed to the robot during the interaction, and whether any other visitors interacted with the robot during the same time, such as when a group of people took turns interacting.

At various points during the study, we approached visitors who had just interacted with the robot to conduct a brief anonymous verbal survey. Initially, we asked participants to rate the robot’s valence, arousal, and mood appropriateness, as well as their entertainment, on a 5-point scale. As the study progressed, we determined a need for additional data in the positive and neutral conditions, resulting in a second survey, which asked about the robot’s valence and arousal, how natural, likable, and entertaining the robot was, the clarity (or enigma) of the robot’s thoughts, and the person’s comfort and mood level with the robot, using a 7-point scale.

5.2.2. Results

We considered the number of seconds each person spent with the robot as a measure of how much that person interacted with it. We considered the following independent variables in our analyses:

- *Mood* refers to the affective condition of the robot: positive, negative, or neutral.
- Based on prior experience with Valerie, we had reason to suspect that the *time of day* would influence people’s interactions with the robot. We grouped times into three categories: morning (before noon), afternoon (noon to 3 pm), and evening (after 3 pm).
- We also considered whether a person interacted with Valerie as part of a *group* of interactors or individually.

Over the course of the experiment, we observed that several of the weeks coincided with events at the University that brought unusually large numbers of visitors past the robot’s booth. Because we observed different interaction patterns during these weeks of high visitor traffic versus weeks of low visitor traffic, we conducted separate analyses for each level. We began with a three-way (affect \times time \times group) analysis of variance (ANOVA) on the time spent with the robot. Given the large number of observations, we chose to remove higher-ordered interactions with F values less than 1 from the analysis. This allowed us in many cases to report on effects

Table 1

Days, number of interactors (I) in each condition, and average interactors per day (I/D), by high and low visitor traffic weeks.

Condition	Low traffic			High traffic		
	Days	I	I/D	Days	I	I/D
Neutral	9	458	50.9	7	483	69.0
Positive	7	410	58.6	8	636	79.5
Negative	8	380	47.5	4	312	78.0
Total	24	1248	52.0	19	1431	75.3

of single variables (i.e. main effects) without the added confusion of interactions between multiple variables. We will not discuss dropped terms. Furthermore, due to space constraints, we focus here only on the influence of the robot's mood (further analyses can be found in [34]).

5.2.2.1. Low visitor traffic weeks. Few visitors came to the University during five of the nine weeks of the study. That is, these weeks had no major events that brought a large number of visitors to the University, so most of the people who interacted with the robot likely had some degree of familiarity with it. The robot was available for 24 days during these weeks, during which time a total of 1248 people interacted with it by typing at least one line of text. The number of interactors in each condition is given in Table 1.

After dropping terms as described above, our final model of seconds included the main effects of mood, time of day, and group, and the interaction of mood and time of day. This model predicted approximately 1.8% of the variance of seconds spent talking with the robot. The interaction between mood and time of day was not significant ($F[4, 1230] = 1.32, n.s.$). We found significant main effects for each of mood ($F[2, 1230] = 2.95, p = 0.05$), time of day ($F[2, 1230] = 2.92, p = 0.05$), and group ($F[1, 1230] = 8.55, p = 0.004$).

The main effect of mood can be explained as follows. People interacted with Valerie for similar times in the neutral and negative conditions (neutral $M = 69.0$, negative $M = 67.7$), but for less time in the positive condition ($M = 57.9$), as can be seen in Fig. 12. Comparisons of mean values (contrast tests) showed that the neutral condition did not differ significantly from both mood conditions ($F[1, 1238] = 0.73, n.s.$). However, people interacted for significantly less time in the positive condition than in the neutral condition ($F[1, 1238] = 3.70, p = 0.05$).

5.2.2.2. High visitor traffic weeks. We classified the remaining four weeks as “high visitor traffic” weeks. During these weeks, unusually high numbers of University visitors passed through the building in which the robot is located, resulting from events such as summer classes for visiting high school students and scheduled visits for incoming undergraduates and their families. The robot was available for 19 days during these weeks, during which time 1431 people interacted with it (Table 1). Significantly more people interacted with the robot per day during the “high visitor traffic” weeks than during the “low” weeks ($F[1, 41] = 19.44, p < .001$), which is consistent with the hypothesis that many people who interacted with the robot during these weeks were visitors who were likely unfamiliar with the robot.

As with low-traffic weeks, we performed an ANOVA on seconds, modeling all main and interaction effects of mood, time of day, and group. The three-way interaction effect did not satisfy the $F < 1$ heuristic for removal discussed earlier, so the model was not changed. This model explained approximately 2.8% of the variance in seconds. The three-way interaction was not significant ($F[4, 1413] = 1.12, n.s.$), nor were the two-way interactions between mood and time of day ($F[4, 1413] = 1.18, n.s.$) or between mood and group ($F[2, 1413] = 0.81, n.s.$). Unlike the “low” weeks, there was a significant interaction effect between group

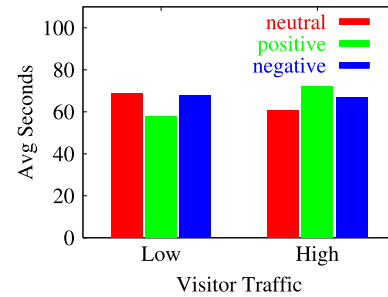


Fig. 12. Average seconds spent with the robot in each mood condition, compared across low and high visitor traffic weeks.

Table 2

Average ratings of the robot's valence and arousal from Survey 1. All variables are on a scale of 1–5. $N = 62$.

Question	Neutral		Positive		Negative	
	M	SD	M	SD	M	SD
Valence	2.73	1.03	3.00	1.15	2.50	0.86
Arousal	2.50	0.88	3.06	1.24	2.45	1.14

and time of day ($F[2, 1413] = 3.23, p = 0.04$). Additionally, there were significant main effects of mood ($F[2, 1413] = 4.18, p = 0.02$), time of day ($F[2, 1413] = 3.35, p = 0.04$), and group ($F[1, 1413] = 15.0, p < 0.001$).

We performed contrast tests to describe the effects of mood. In contrast to the “low” weeks, people in the neutral condition interacted for a significantly shorter period of time than in either emotional condition (neutral $M = 60.1$, positive $M = 72.4$, negative $M = 66.7$; $F[1, 1413] = 4.53, p = 0.03$). The number of seconds spent interacting in the negative and positive conditions were not significantly different ($F[1, 1413] = 1.64, n.s.$).

5.2.2.3. Survey 1. The two surveys were designed to measure visitors' subjective experiences of the robot. Over two weeks, we collected 62 responses to the first survey: 16 on positive days, 22 on negative days, and 24 on neutral days. This survey measured participant's ratings of the robot's valence and arousal, appropriateness of mood to what was said (mood congruency), and their entertainment. Initially, we placed the surveys on the robot's booth with a sign requesting interactors to complete the survey. As this collection method proved insufficient, the remainder of the surveys were collected by hand. The “self-survey” condition was included as an independent variable in the following analyses.

Valence and arousal were correlated ($r = 0.43, p < 0.001$), but no other measures were significantly correlated (at the $p < 0.01$ level), indicating that each question measured a different aspect of the visitor's experience. When analyzed individually, none of the four questions (i.e. valence, arousal, mood congruency, and entertainment) differed significantly between the three mood conditions. However, by analyzing people's ratings of the robot's valence and arousal with a multivariate analysis of variance (MANOVA), we found that people were able to differentiate between the mood conditions ($F[2, 58] = 2.39, p = 0.10$). The positive condition had the highest valence/arousal ratings, negative had the lowest, and neutral fell in between, as shown in Table 2. The differences were marginal, but this may be a result of the small sample size or the granularity of the scales used.

Neither appropriateness nor entertainment differed significantly across the mood conditions (appropriateness $F[2, 58] = 1.68, n.s.$; entertainment $F[2, 58] = 1.62, n.s.$). Overall, respondents responded near the scale's midpoint for both the appropriateness of the robot's mood ($M = 3.63, SD = 1.27$) and their entertainment ($M = 3.19, SD = 1.27$). The presence of the experimenter had an effect only on the rating of appropriateness, with

Table 3Results from Survey 2, by robot condition. Variables are on a scale of 1–7. $N = 61$.

Question	Neutral		Positive		F	p
	M	SD	M	SD		
Robot valence	3.91	1.06	4.27	1.43	1.08	0.30
Robot arousal	3.00	0.93	3.76	1.23	6.57	0.01
Natural	3.46	1.14	4.08	1.28	3.75	0.06
Likable	4.58	1.18	4.30	1.29	0.77	0.38
Entertainment	4.46	1.18	3.97	1.54	1.73	0.19
Enigmatic	5.08	1.21	4.22	1.57	5.29	0.03
Comfort	5.42	1.28	5.08	1.26	1.02	0.32
Person valence	5.33	1.24	4.49	1.33	6.25	0.02

individuals responding on their own rating the robot's mood as less appropriate (self-survey $M = 3.00$, face-to-face $M = 3.78$; $F[2, 58] = 6.99$, $p = 0.01$).

5.2.2.4. Survey 2. We collected 61 responses to the second survey: 37 on positive days, and 24 on neutral days. The survey was designed to investigate the low interaction times in the positive condition during low-traffic weeks, and so no responses were collected on negative days. As discussed, this survey contained questions regarding both the robot's and the respondent's states. Results for each question are shown in Table 3.

As with the first survey, we computed a MANOVA on the robot's valence and arousal. Respondents rated the robot's valence/arousal as significantly higher in the positive condition than in the neutral condition ($F[1, 59] = 4.22$, $p = 0.04$). They found the positive robot significantly less enigmatic and slightly more natural than the neutral robot. There were no significant differences in ratings of the robot's likability, entertainment, or participant's comfort.

5.2.3. Discussion

Both surveys showed that people who interacted with the robot could distinguish between the robot's moods, and the robot's mood did have an effect on people's interactions with it. However, this effect varied across high and low visitor traffic weeks. The differences were fairly large—on the order of ± 10 s, while interactions averaged only about a minute. As we expected, more people interacted with the positive robot, and fewer with the negative robot, during the low visitor traffic weeks. Though we expected that people would interact longest with the positive robot and shortest with the negative, we found that people interacted for significantly less time with the positive robot during those weeks. We believe that this finding can be explained by relying on the theory of common ground [35]. Past studies have shown that people's perception of common ground with robots results in shorter, more "efficient" speech [36]. Smiling carries a certain amount of conversational content, so people may not have felt that they needed to interact as long with the positive robot to come to the same level of understanding as when the robot did not smile. This theory is supported by the results of Survey 2; participants found the positive robot to be significantly less enigmatic—that is, easier to understand—than the neutral robot. Additionally, participants found the positive robot more natural than the neutral robot, which could also contribute to a sense of common ground.

During high-traffic weeks, in contrast, visitors interacted the least in the neutral condition and the most in the positive condition. One possible explanation for this result may be that the moody robot appeared more interesting; perhaps any form of affect in a robot may be compelling enough to sustain interactions. Though we did not analyze the common ground theory during the high-traffic weeks, we believe that people with little experience with the robot may have required more time to form common ground with the robot. That is, because visitors to the University may have lacked prior knowledge of the robot, they may not have

noticed the robot's smile (or frown) as a *change* from prior behavior, and thus not have felt stronger common ground with the smiling robot. Furthermore, interactions during the high-traffic weeks were likely biased due to a novelty effect, while visitors during low-traffic weeks may have been more used to the robot's presence.

A potential shortcoming of our analysis is that we do not currently have any means of automatically identifying whether any particular person has previously interacted with the robot or even been familiar with the robot prior to his or her interaction during this study, and so we cannot be certain that the differences between high and low visitor traffic weeks are primarily due to the visitors. However, we do know that the weeks we considered "high visitor traffic" did have an unusually high number of visitors to the University. Those same weeks also had more interactors per day and different interaction patterns than usual. Thus we feel that our assumptions regarding people's familiarity with the robot are valid.

5.3. Summary

We have run two studies that have helped validate the emotion and mood components of our affect model. The first study found that people can identify the robot's expressions of happiness, sadness, anger, and disgust and, furthermore, that people can distinguish between different intensity levels of each emotion. This study also showed that people do not necessarily assign the same valence and arousal ratings to the same emotions on different robotic faces (i.e. Valerie versus Tank), which experimenters need to consider when designing studies on different robots. The second study showed significant differences in how people interact with a robot based on its apparent mood. In particular, the study showed that people tend to change their interaction patterns with the robot when it displays different moods, but that exactly how people's patterns change depends on whether they have prior familiarity with the robot. In general, these findings are consistent with the study of common ground in human–human interaction.

Because not many people choose to identify themselves to the robot (using the cardreader), we have not yet validated the attitude portion of the affective model. Such research will be conducted once the robot is better able to identify visitors. Since we have shown that people interact with the robot differently based on their prior familiarity with the robot, we believe that better incorporating attitudes toward individuals will also enhance people's interactions with the robot.

6. Conclusion

Humans increasingly need to communicate with robots, and we are working toward making human–robot interaction be smooth and natural. This paper has presented a generative model of affect that accounts for emotions, moods, and attitudes, including the interactions between them. The model tries to mimic human behavior, particularly with regard to long-term affective responses. We have presented our implementation of this model on the Roboceptionist, along with findings from studies that tested the emotion and mood components of the model.

Our research shows that people are able to understand the robot's expressions of emotion and that only slight changes in a robot's expressions—and no changes in the structure of the interactions—influence how people perceive and interact with the robot. Different expressions create different expectations of the robot, with many people avoiding the negative robot or feeling more common ground with the positive robot. Furthermore, we found that people with different levels of familiarity with the robot react to the robot's mood in different ways. This shows that a social robot needs to remember people who have interacted with it, and interact differently with those people than with newcomers. The

robot should utilize its emotional expressions differently based on how much common ground it shares with a person, which is addressed by the attitude portion of our affective model.

We have had this model running on the Roboceptionist system—first on Valerie and currently on Tank—since August 2005. We have been working to improve the robot's person identification system in order to fully test the robot's attitudes toward different people. In addition, such identification is necessary to understand how the affective model as a whole influences human-robot interaction. Furthermore, we hope to extend the testing of this model to other robots, as we have found that people do not view similar expressions on different faces to be identical, even when displayed on the same physical robot. We believe that this model will prove useful for a wide variety of social robots, as we believe that a robot's moods and emotions will greatly improve its ability to communicate with people.

References

- [1] R. Gockley, A. Bruce, J. Forlizzi, M. Michalowski, A. Mundell, S. Rosenthal, B. Sellner, R. Simmons, K. Snipes, A.C. Schultz, J. Wang, Designing robots for long-term social interaction, in: Proceedings of IROS 2005, Edmonton, Alberta, 2005.
- [2] B. Parkinson, Emotions are social, *The British Psychological Society* 87 (1996) 663–683.
- [3] H.H. Clark, S.E. Brennan, Grounding in communication, in: *Perspectives on Socially Shared Cognition*, APA, 1991, pp. 127–149.
- [4] J.P. Forgas, Feeling and speaking: Mood effects on verbal communication strategies, *Personality and Social Psychology Bulletin* 25 (7) (1999) 850–863.
- [5] B. Wild, M. Erb, M. Bartels, Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences, *Psychiatry Research* 102 (2) (2001) 109–124.
- [6] N.H. Frijda, Emotion experience, *Cognition and Emotion* 19 (4) (2005) 473–497.
- [7] E.A. Butler, B. Egloff, F.H. Wilhelm, N.C. Smith, E.A. Erickson, J.J. Gross, The social consequences of expressive suppression, *Emotion* 3 (1) (2003) 48–67.
- [8] B. Reeves, C. Nass, *The Media Equation*, CSLI Publications, Cambridge, 1996.
- [9] I. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, A. Soto, An affective mobile robot educator with a full-time job, *Artificial Intelligence* 114 (1–2) (1999) 95–124.
- [10] M. Montemerlo, J. Pineau, N. Roy, S. Thrun, V. Verma, Experiences with a mobile robotic guide for the elderly, in: *Proceedings of the National Conference of Artificial Intelligence (AAAI 02)*, Edmonton, AB, 2002.
- [11] C. Breazeal, Emotion and sociable humanoid robots, *International Journal of Human Computer Studies* 59 (2003) 119–155.
- [12] A. Bruce, I. Nourbakhsh, R. Simmons, The role of expressiveness and attention in human-robot interaction, in: *Proceedings of the IEEE International Conference on Robotics and Automation, ICRA, 2002*.
- [13] C. Elliott, *The affective reasoner: A process model of emotions in a multi-agent system*, Ph.D. Thesis, Northwestern University, 1992.
- [14] J. Cassell, T. Bickmore, L. Campbell, H. Vilhjálmsdóttir, H. Yan, Human conversation as a system framework: Designing embodied conversational agents, in: Cassell et al. [15], pp. 29–63.
- [15] J. Cassell, J. Sullivan, S. Churchill (Eds.), *Embodied Conversational Agents*, MIT Press, 2000.
- [16] A. Paiva, J. Dias, D. Sobral, R. Aylett, S. Woods, L. Halle, C. Zoll, Learning by feeling: Evoking empathy with synthetic characters, *Applied Artificial Intelligence* 19 (3–4) (2005) 235–266.
- [17] A. Ortony, G.L. Clore, A. Collins, *The Cognitive Structure of Emotions*, Cambridge University Press, 1988.
- [18] M. El Jed, N. Pallamin, J. Dugdale, B. Pavard, Modelling character emotion in an interactive virtual environment, in: *Proceedings of AISB 2004 Symposium on Motion, Emotion and Cognition*, Leeds, UK, 2004.
- [19] L. Moshkina, R.C. Arkin, On TAMEing robots, in: *IEEE International Conference on Systems, Man and Cybernetics, 2003*.
- [20] K.R. Scherer, Psychological models of emotion, in: J.C. Borod (Ed.), *The Neuropsychology of Emotion*, Oxford University Press, Oxford, 2000, pp. 137–162.
- [21] C.E. Izard, Innate and universal facial expressions: Evidence from developmental and cross-cultural research, *Psychological Bulletin* 115 (2) (1994) 288–299.
- [22] J.A. Russell, J.-A. Bachorowski, J.-M. Fernández-Dols, Facial and vocal expressions of emotion, *Annual Review of Psychology* 54 (2003) 329–349.
- [23] P. Ekman (Ed.), *Emotion in the Human Face*, 2nd ed., University of Cambridge Press, Cambridge, England, 1982.
- [24] J. Russell, Core affect and the psychological construction of emotion, *Psychological Review* 110 (1) (2003) 145–172.
- [25] P. Ekman, Moods, emotions, and traits, in: P. Ekman, R.J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions*, Oxford University Press, New York, 1994, pp. 56–58.
- [26] R. Lazarus, The stable and the unstable in emotion, in: P. Ekman, R.J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions*, Oxford University Press, New York, 1994, pp. 79–85.
- [27] J. Neumann, B. Seibt, F. Strack, The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing, *Cognition and Emotion* 15 (6) (2001) 725–747.
- [28] N. Bolger, A. DeLongis, R.C. Kessler, E.A. Schilling, Effects of daily stress on negative mood, *Journal of Personality and Social Psychology* 57 (5) (1989) 808–818.
- [29] K.S. Rook, Emotional health and positive versus negative social exchanges: A daily diary analysis, *Applied Developmental Sciences* 5 (2) (2001) 86–97.
- [30] S.J. Breckler, E.C. Wiggins, On defining attitude and attitude theory: Once more with feeling, in: A.R. Pratkanis, S.J. Breckler, A.G. Greenwald (Eds.), *Attitude Structure and Function*, Lawrence Erlbaum Associates, New Jersey, 1989, pp. 407–427.
- [31] H.L. Wagner, J. Smith, Facial expression in the presence of friends and strangers, *Journal of Nonverbal Behavior* 15 (4) (1991) 201–214.
- [32] C. Bartneck, Integrating the OCC model of emotions in embodied characters, in: *Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenge, 2002*.
- [33] C.L. Rusting, Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks, *Psychological Bulletin* 124 (2) (1998) 165–196.
- [34] R. Gockley, J. Forlizzi, R. Simmons, Interactions with a moody robot, in: *Proceedings of Human-Robot Interaction*, Salt Lake City, Utah, 2006.
- [35] H.H. Clark, *Using Language*, Cambridge University Press, Cambridge, 1996.
- [36] A. Powers, A. Kramer, S. Lim, J. Kuo, S. lai Lee, S. Kiesler, Common ground in dialogue with a gendered humanoid robot, in: *Proceedings of RO-MAN 2005*, Nashville, TN, 2005.
- [37] R. Gockley, R. Simmons, J. Forlizzi, Modeling affect in socially interactive robots, in: *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Hatfield, United Kingdom, 2006.



Rachel Kirby is a Ph.D. candidate in the Robotics Institute at Carnegie Mellon University. Widely, her research is in the area of human-robot social interaction. In particular, her thesis work focuses on the issue of social spatial navigation, allowing robots to navigate safely and comfortably with and around people. She received her M.S. in Robotics in 2005 and her B.S. in Computer Science and Philosophy in 2003, both from Carnegie Mellon University.



Jodi Forlizzi is an Associate Professor of Design and Human-Computer Interaction at Carnegie Mellon University, in Pittsburgh, PA. She is an interaction designer contributing to design theory and practice. Her theoretical research examines theories of experience, emotion, and social product use as they relate to interaction design. Other research and practice centers on notification systems ranging from peripheral displays to embodied robots, with a special focus on the social behavior evoked by these systems.



Reid Simmons is a Research Professor in the School of Computer Science at Carnegie Mellon University. He earned his B.A. degree in 1979 in Computer Science from SUNY at Buffalo, and his M.S. and Ph.D. degrees from MIT in 1983 and 1988, respectively, in the field of Artificial Intelligence. His thesis work focused on the combination of associational and causal reasoning for planning and interpretation tasks. The research analyzed the relationships between different aspects of expertise and developed a domain-independent theory of debugging faulty plans. Since coming to Carnegie Mellon in 1988, Dr. Simmons' research has focused on developing self-reliant robots that can autonomously operate over extended periods of time in unknown, unstructured environments. This work involves issues of robot control architectures that combine deliberative and reactive control, probabilistic planning and reasoning, monitoring and fault detection, and robust indoor and outdoor navigation. More recently, Dr. Simmons has focused on the areas of coordination of multiple heterogeneous robots, human-robot social interaction (including the robots Grace and the Roboceptionist), and formal verification of autonomous systems. Over the years, he has been involved in the development of over a dozen autonomous robots.