Relational Agents: A Model and Implementation of Building User Trust

Timothy Bickmore, Justine Cassell
MIT Media Lab
20 Ames St., E15-320
Cambridge, MA 02139 USA
+1 617 253 7368
{bickmore, justine}@media.mit.edu

ABSTRACT
Building trust with users is crucial in a wide range of applications, such as financial transactions, and some minimal degree of trust is required in all applications to even initiate and maintain an interaction with a user. Humans use a variety of relational conversational strategies, including small talk, to establish trusting relationships with each other. We argue that such strategies can also be used by interface agents, and that embodied conversational agents are ideally suited for this task given the myriad cues available to them for signaling trustworthiness. We describe a model of social dialogue, an implementation in an embodied conversation agent, and an experiment in which social dialogue was demonstrated to have an effect on trust, for users with a disposition to be extroverts.

Keywords
Embodied conversational agent, trust, social interface, natural language, small talk, personality.

INTRODUCTION
Humans use a variety of strategies to proactively establish and maintain social relationships with each other. Building rapport and common ground through small talk, intimacy through self-disclosure, credibility through the use of expert's jargon, social networks through gossip, and "face" through politeness are all examples of this phenomenon. These relational strategies are important not just in purely social settings, but are also crucial to the establishment and maintenance of any collaborative relationship.

Computer interface agents may also profitably use relational strategies such as these if they are to function successfully in roles which require users to interact with them for more than a few minutes, or in which we expect users to take them seriously enough to discuss their medical problems or give out their credit card numbers. Agents of this sort must be able to establish social relationships with users in order to engage their trust which, in turn, eases cooperation.

SIGCHI'01, March 31-April 4, 2001, Seattle, WA, USA.

Existing "social" interface agents (e.g., Microsoft "Bob" or the Paper Clip) achieve their social effects by attempting to draw the user into what is billed as a social interaction; essentially a passive strategy for relationship building. What these systems lack are explicit behaviors, protocols and strategies for building, maintaining or changing a relationship with the user, something humans have a large repertoire of techniques for. Further, these systems make poor use of the primary modality humans use to establish and maintain relationships, namely language.

Embodied Conversational Agents (ECAs) are particularly well suited to the task of relationship building. ECAs are anthropomorphic interface agents which engage a user in real-time dialogue, using speech, gesture, gaze, and other verbal and nonverbal channels to emulate the experience of human face-to-face interaction. The nonverbal channels are important for conveying information, and for regulating the flow of the conversation. These nonverbal channels are also especially crucial for relational conversation, since they can be used to provide such social cues as attentiveness, positive affect, and liking and attraction, and to mark shifts into and out of relational activities.

In this paper we will discuss a model of social dialogue for building user trust: we will talk about the conversational strategies that comprise the model, and one kind of talk -- small talk—that executes those strategies. Finally, we will describe an evaluation of our approach where users interacted with one of two embodied conversational agents, and we evaluated their trust in the interaction. We concentrate on the relational notion of trust because it is essential for all kinds of interpersonal interactions, and crucially important for certain types of human-computer interactions [12]. Trust between humans involves credibility, believing one another, confidence in another's judgments, and belief that another's actions fit our own schemata of how to act. Trust is a prerequisite for actions involving another agent in which one may suffer physical, financial or psychological harm (e.g., financial transactions, or disclosing personal information [30]).

Related Work in Relational Agents

In a series of studies in the "Computers As Social Actors" paradigm, researchers have demonstrated the possibility of manipulating the user's relationship with a computer using a wide range of behaviors. Reeves & Nass demonstrated that users like computers more when the computer flatters them [21]. Morkes, Kernal and Nass demonstrated that computer agents which use humor are rated as more likable, competent and cooperative [18]. Moon demonstrated that a computer which uses a strategy of reciprocal, deepening self-disclosure in its (text-based) conversation with the user will cause the user to rate it as more attractive, divulge more intimate information, and become more likely to buy a product from the computer [19].

Of course the social influence strategies of relational agents may not be equally effective across all types of users. Several studies have shown that users react differentially to social agents based on their own personality and other dispositional traits. For example, Reeves and Nass have shown that users like agents that match their own personality (on the introversion/extraversion dimension) more than those which do not, regardless of whether the personality is portrayed through text or speech [21] [20]. Resnick and Lammers showed that in order to change user behavior via corrective error messages, the messages should have different degrees of "humanness" depending on whether the user has high or low self-esteem ("computer-ese" messages should be used with low self-esteem users, while "human-like" messages should be used with high-esteem users) [22]. Rickenberg and Reeves showed that different types of animated agents differentially affected the anxiety level of users as a function of whether users tended towards internal or external locus of control [24].

Embodied Conversational Agents

Work on the development of ECAs, as a distinct field of development, is best summarized in [8]. In addition to REA [6] (described below), some of the other major ECA systems developed to date are Steve [23], the DFKI Persona [1], Olga [3], Gandalf [29], and pedagogical agents developed by Lester, et al, [15, 16]. There are also a growing number of commercial ECAs, such as those developed by Extempo, Headpedal, and Artificial Life, and the Ananova newscaster developed by Ananova, Ltd. These systems vary greatly in their linguistic capabilities, input modalities (most are mouse/text/speech input only), and task domains, but all share the common feature that they attempt to engage the user in natural, full-bodied (in some sense) conversation. Although these systems hold out the promise of increased engagement and effectiveness, evaluations of their use in domains from learning and training to entertainment and communication have not proved their worth. Dehn and van Mulken [11], specifically examining evaluations of recent animated interface agents, conclude that the benefits of these systems are arguable in terms of user performance, engagement with the system, or even attributions of intelligence. However, they point out that virtually none of the systems evaluated exploited the affordances of human bodies: this design paradigm "can only be expected to improve human–computer interaction if it shows some behavior that is functional with regard to the system's aim." In light of these results, we have designed an embodied conversational agent that is based on a model of social dialogue for building user trust and diminishing interpersonal distance, and that is implemented in a domain in which exactly these abilities are key.

A MODEL OF SOCIAL DIALOGUE FOR USER TRUST

Interpersonal relationships can be measured along many dimensions, including intimacy, solidarity, closeness, familiarity, and affiliation [26]. Since we are primarily interested in dimensions that have an effect on trust and that can be employed to formulate a communicative strategy, we base our user-computer social linguistic model on three dimensions of the 'interpersonal relations in conversation' model developed by Svennevig [28]. In what follows, we describe these three dimensions, and some strategies for affecting them, from Svennevig's own model, and then we lay out our own extensions to the model.

The first dimension of Svennevig's relational model is labeled *familiarity*, and accounts for the way in which relationships develop through the reciprocal exchange of information, beginning with relatively non-intimate topics and gradually progressing to more personal and private topics. The growth of a relationship can be represented in both the breadth and depth of information disclosed.

Two other dimensions of Svennevig's relational model--*power* and *solidarity*--have been dealt with both in social psychology and linguistics. Power is the ability of one interactant to control the behavior of the other. Solidarity is defined as "like-mindedness" or having similar behavior dispositions (e.g., similar religion, profession, gender, etc.). There is a correlation between frequency of contact and solidarity, but it is not necessarily a causal relation [4, 5].

Although trust is also an essential part of human social relationships, and is often established through linguistic means, following Svennevig our model does not include trust as one of the dimensions, since it can be better viewed as a function or outcome of the above attributes, and not a dimension to be modeled independently. Trust can be defined as "people's abstract positive expectations that they can count on partners to care for them and be responsive to their needs, now and in the future," and one model of the development of trust describes it as "a process of uncertainty reduction, the ultimate goal of which is to reinforce assumptions about a partner's dependability with actual evidence from the partner's behavior" [2]. Thus, trust is predicated on solidarity and familiarity, but also includes information about specific trusting behaviors. Note that this formulation differs from recent work on trust in the computational community in that work on trust in e-

commerce or among agents often relies on transaction characteristics rather than interpersonal characteristics.

Conversational Strategies for Changing Interpersonal Relationships

Our objective is to build an ECA that knows how to win people's trust and that goes about the process using relational conversational strategies. This requires a model of trust that is broken down into the goals to be achieved and the conversational strategies for achieving them, as well as the ways of generating those conversational strategies and putting them into practice. In this section we explain two broad categories of conversational strategy that play a role in achieving increased trust -- facework, and establishing common ground. We then turn to how these strategies can be generated and put into practice in small talk generated by an ECA.

In Goffman's approach to social interaction, he defined an interactant's "line" as the patterns of action by which individuals in an interaction present an image of themselves and the situation [13]. The notion of "face" is "the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact". Interactants maintain face by having their line accepted and acknowledged. Events which are incompatible with their line are "face threats" and are mitigated by various corrective measures if they are not to lose face. In short, events which are incompatible with how we wish others to see us, are called "face threats", and we try to avoid them, and to mitigate their effect if they are unavoidable.

Brown and Levinson extended Goffman's notion of face in their theory of politeness forms in language [4]. They characterized the degree of face threat of a given speech act as a function of power, social distance, and the intrinsic threat (imposition) imposed by the speech act.

Based on our own analysis of social dialogue in service encounters, we have further extended Brown and Levinson's model of face threats. Given the relational model presented above, the introduction of conversational topics which are at a significantly "deeper" level of familiarity than is expected relative to the existent relationship and activity are seen as a face threat. For example, if a stranger on the street asked you how much money you had in your bank account, you would likely perceive this as a threat to your face.

How can speakers change these dimensions of trust? One strategy for effecting changes to the familiarity dimension of the relationship model is for the speaker to disclose information about him/herself and induce the listener to do the same. Another way of changing the dimensions of trust in conversation is to engage in small talk.

Small Talk: Putting Trust-Elicitation into Practice

Small talk can be taken as any talk in which interpersonal goals are emphasized and task goals are either non-existent or de-emphasized. Within task-oriented encounters, small talk can help humans or agents to achieve their goals by "greasing the wheels" of task talk. It can serve a transitional function, providing a ritualized way for people to move into conversation in what may be an otherwise awkward situation [14]. Small talk can also serve an exploratory function by providing a conventional mechanism for people to establish the "communal common ground" [10] of another human or a computational system. Small talk can build solidarity through a ritual of showing agreement with and appreciation of the conversational partner's utterances [18], [9, 25]. Finally, people and agents can use small talk to establish expertise, by relating stories of past successful problem-solving behavior, and to obtain information about the other that can be used indirectly to help achieve task goals (e.g., that the user drives a minivan increases the probability that the person has children).

Small talk can be used to address the face needs of interlocutors. In small talk, interlocutors take turns showing agreement with and appreciation of the contributions of the speaker, and in so doing enhance each other's face [9, 25]. This builds solidarity among the interlocutors by demonstrating their "like mindedness". Small talk can also be used in social situations as a prelude to other more personal kinds of talk once the interlocutors decide that they want to move on to the next stage of their relationship.

Thus, small talk implements the conversational strategies listed above in order to build trust (see Figure 1). It acts on a peer relationship among interlocutors, and thus may help to side-step any power imbalance. It allows them to establish common ground and increase their familiarity. It increases solidarity through mutual acknowledgement. In fact, interaction rituals such as these also fit into the uncertainty reduction model of trust, in which individuals incrementally reinforce their assumptions about the partner's dependability with actual evidence from his/her behavior [2]. The natural progression of a conversation

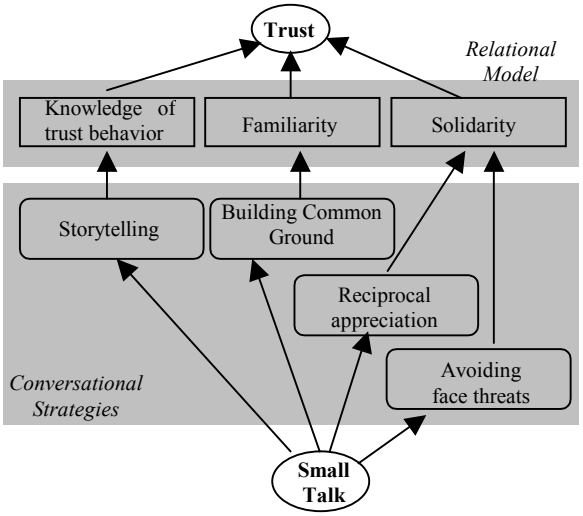

Figure 1. Influence of Small Talk on Trust

between strangers from greetings, through small talk, into more substantive topics can be seen as a process in which they iteratively "test the water" to determine if they want to continue deepening the relationship.

AN IMPLEMENTATION: SMALL TALK IN REA

REA is a real-time, multimodal, life-sized ECA, and her design is based on the FMBT model [6, 7]. REA has a fully articulated graphical body, can sense the user passively through cameras and audio input, and is capable of speech with intonation, facial display, and hand gesture. REA is displayed on a large projection screen, in front of which the user stands (see Figure 2). Two cameras mounted on top of the screen track the user's head and hand positions, while a microphone captures speech input. A single SGI Octane computer runs the graphics and conversation engine of Rea, while several other computers manage the speech recognition and generation, and image processing.

Figure 2. User interacting with Rea

Rea simultaneously processes the organization of conversation and its content. When the user makes cues typically associated with turn taking behavior such as gesturing, Rea allows herself to be interrupted, and then takes the turn again when she is able. An incremental natural language generation engine [27], extended to synthesize redundant and complementary conversational hand gestures, generates Rea's responses. REA is an acronym for "Real Estate Agent", and within this domain we model the initial interview with a prospective buyer. Real estate sales was selected specifically for the opportunity to explore a task domain in which a significant amount of social dialogue normally occurs.

Implementing Relational Strategies in REA

Within initial interactions between professionals and their clients, small talk is often used to build trust and solidarity. This is especially important in real estate, where the stakes are high and the buyer-agent relationship must continue for several weeks or months until a transaction is closed.

For the purpose of trust elicitation and small talk, we have constructed a new kind of discourse planner that can interleave small talk and task talk during the initial buyer interview, based on the model outlined above. Given that many of the goals in a relational conversational strategy are non-discrete (e.g., minimize face threat), and that trade-offs among multiple goals have to be achieved at any given time, we have moved away from static world discourse planning, and use an activation network-based approach based on [17]. This architecture can transition smoothly from deliberative, planned behavior to opportunistic, reactive behavior, and can pursue multiple, non-discrete goals. In our implementation each node in the network represents a conversational move that REA can make.

During task talk, REA asks questions about users' buying preferences, such as the number of bedrooms they need. During small talk, REA can talk about the weather, events and objects in her shared physical context with the user, or she can tell stories about the lab, herself, or real estate.

REA's contributions to the conversation are planned in order to minimize face threat and maximize trust, while pursuing her task goals in the most efficient manner possible. That is, Rea attempts to determine the face threat of her next conversational move, assesses the solidarity and familiarity which she currently holds with the user, and judges which topics will seem most relevant and least intrusive to users. As a function of these factors, Rea chooses whether or not to engage in small talk, and what kind of small talk to choose. The selection of which move should be pursued by REA at any given time is thus a non-discrete function of the following factors:

- Closeness -- Rea continually assesses her "interpersonal" closeness with the user, which is a composite representing depth of familiarity and solidarity, modeled as a scalar quantity. Each conversational topic has a pre-defined, pre-requisite closeness that must be achieved before Rea can introduce the topic. Given this, the system can plan to perform small talk in order to "grease the tracks" for task talk, especially about sensitive topics like finance.
- Topic -- Rea keeps track of the current and past conversational topics. Conversational moves which stay within topic are given preference over those which do not. In addition, Rea can plan to execute a sequence of moves which gradually transition the topic from its current state to one that Rea wants to talk about (e.g., from talk about the weather, to talk about Boston weather, to talk about Boston real estate).
- Relevance -- Rea maintains a list of topics that she thinks the user knows about, and the discourse planner prefers moves which involve topics in this list. The list is initialized to things that anyone talking to Rea would know about--such as the weather outside, Cambridge, MIT, or the laboratory that Rea lives in.
- Task goals -- Rea has a list of prioritized goals to find out about the user's housing needs in the initial interview. Conversational moves which directly work towards satisfying these goals (such as asking interview questions) are preferred.
- Logical preconditions -- Conversational moves have logical preconditions (e.g., it makes no sense for Rea to ask users their major until she has established that they are students), and are not selected for execution until all of their preconditions are satisfied.

One advantage of the activation network approach is that by simply adjusting a few gains we can make REA more or less coherent, more or less attentive to closeness constraints, more or less task-oriented, or more or less deliberative (vs. reactive) in her linguistic behavior.

In the current implementation, the dialogue is entirely REA-initiated, and user responses are recognized via a speaker-independent, grammar-based, continuous speech recognizer (IBM ViaVoice). The active grammar fragment is specified by the current conversational move, and for responses to many Rea small talk moves the content of the user's speech is ignored; only the fact that the person responded at all is enough to advance the dialogue.

At each step in the conversation in which Rea has the floor (as tracked by a conversational state machine), the discourse planner is consulted for the next conversational move to initiate. At this point, activation values are incrementally propagated through the network (following [17]) until a move is selected whose preconditions are satisfied and whose activation value is over a specified threshold.

Shifts between small talk moves and task moves are marked by conventional contextualization cues--discourse markers and beat gestures. Discourse markers include "so" on the first small talk to task talk transition, "anyway" on resumption of task talk from small talk, and "you know" on transition to small talk from task talk [10].

Within this framework, Rea decides to do small talk whenever closeness with the user needs to be increased (e.g., before a task query can be asked), or the topic needs to be moved little-by-little to a desired topic and small talk contributions exist which can facilitate this. The activation energy from the user relevance condition described above leads to Rea starting small talk with topics that are known to be in the shared environment with the user.

Example Interactions

An interview between REA and a user typically proceeds as shown in the following dialogue. (User responses are only shown in positions in which they affect the selection of subsequent moves)

1. That microphone is terrible, I hate using those things.
2. Sorry about my voice, this is some engineer's idea of natural sounding.
3. Are you one of our sponsors? *User: Yes*
4. Were you at our last sponsor meetings?
5. I got so exhausted at the last sponsor meeting I think I was starting to lose my voice by the end.
6. So, where would you like to live?
7. How many bedrooms do you need?
8. Do you need access to the subway?
9. Is one bath enough?
10. You know, Boston is certainly more expensive than it used to be.
11. Anyway, what can you afford?
12. What kind of down payment can you make?
13. Let me see what I have available.

Dialogue 1. "Small Talk REA"

In this example, REA opens with small talk moves regarding things in her shared physical environment with the user (1-2). She then proceeds to small talk related to sponsors (after establishing that the user is a sponsor). After a few turns, enough closeness has been established (simply by doing small talk) that REA can move into task talk (6-9). However, before bringing up the potentially face-threatening topic of finance REA decides that additional closeness needs to be established, and moves back into small talk (10). This small talk move increases closeness and shifts the topic to finance, enabling REA to ask how much the user is able to afford (11-12).

If REA's adherence to closeness preconditions is reduced, by decreasing the contributions of these preconditions to the activation of conversational moves, this results in her engaging in less small talk and being more task goal oriented. If everything else is held constant (relative to the prior example) the following dialogue is produced.

1. So, where would you like to live?
2. What can you afford?
3. What kind of down payment can you make?
4. How many bedrooms do you need?
5. Do you need access to the subway?
6. Is one bath enough?
7. Let me see what I have available.

Dialogue 2. "Task-only REA"

In this example, REA performs no small talk and sequences the task questions in strictly decreasing order of priority.

EVALUATION

To evaluate whether an ECA's social dialogue can actually build trust and solidarity with users, we conducted an empirical study in which subjects were interviewed by Rea about their housing needs, shown two "virtual" apartments, and then asked to submit a bid on one of them. In the experiment, Rea was controlled by a human wizard and followed scripts identical to the output of the planner (but faster, and not dependent on speech recognition).

Our hypotheses follow from the literature on small talk and on trust among humans. We expected subjects who interact with a version of REA which used small talk to trust her more, like her more, think she was more credible, and feel that they understand each other more. We also expected these users to think the interaction was more natural, satisfying, and successful. Finally, we expected users to be willing to pay REA more for an apartment when she used small talk, given the hypothesized increase in trust.

Experimental Methods

The study was a between subjects design with subjects randomly assigned either to a version of REA which used only task-oriented dialogue (TASK condition) or to an identical version which also included the social dialogue (SMALLTALK condition).

Subjects. 31 people participated in the experiment (58% male and 42% female). Subjects were primarily students,

were recruited through ads on several college campuses, and were compensated for their participation.

Apparatus. An experiment room with one entire wall as a rear-projection screen allowed Rea to appear life-sized on the screen, in front of the 3D virtual apartments she showed. Rea's synthetic voice was played through two speakers on the floor in front of the screen. Two video cameras and an omnidirectional microphone enabled recording of the subject's verbal and nonverbal behavior during the experiment.

The wizard sat behind the rear projection screen and controlled REA's responses and sequencing through the interaction script via a computer. The script included verbal and nonverbal behavior specifications for REA and embedded commands describing when different rooms in the virtual apartments should be shown. Three pieces of information obtained from the user were entered into the control system: the city the subject wanted to live in; the number of bedrooms s/he wanted; and how much s/he was willing to spend. The first apartment shown had twice as many bedrooms as the subject requested and cost twice as much as s/he could afford (subjects were told the price was "firm"). The second apartment shown had the exact number of bedrooms requested, but cost 50% more than the subject could afford (but this time the subject was told that the price was "negotiable"). The scripts for the TASK and SMALLTALK condition were identical, except that the SMALLTALK script had additional small talk utterances, similar to those shown in Dialogue 1, above. The script governing the dialogue from the showing of the second apartment through the end of the interaction was identical in both conditions.

Procedure. Subjects were told that they would be interacting with Rea, who played the role of a real estate agent and could show them apartments she had for rent. They were told to play the role of someone looking for an apartment in the Boston area, and to stand in front of Rea and talk to her "just like you would to another person".

Subjects were shown a brief (one minute) video of REA on a small monitor, giving additional instructions regarding her speech recognition software. The purpose of this was both to reduce the "novelty effect" when REA first appeared on the big projection screen, and to ensure the deception (use of a wizard) was effective. Subjects then interacted with Rea, after which they were asked to fill out a questionnaire.

Manipulation check. Three questions concerning the amount of small talk used by REA were included on the questionnaire, for manipulation checks. There was a significant difference ($F(1,44)=11.2$; $p< .002$) such that users believed that REA got down to business more quickly in the task-only condition than in the small talk condition.

Measures.

Trust was measured by a standardized trust scale taken from [30] (Cronbach's alpha = .88 as measured in [20]).

Evaluation of the interaction was measured as follows. REA's *informedness, knowledgability, credibility, expertise, knowledge of the user,* user's *liking of REA, knowledge of REA, desire to work with REA again, and interest in the interaction*, and *naturalness, satisfaction, engagingness,* and *success of the interaction* were measured by single items on nine-point Likert scales.

Amount Willing to Pay: During the interview, Rea asked subjects how much they were able to pay for an apartment; subjects' responses were entered as $X per month. REA then offered the second apartment for $Y (where Y = 1.5 X), and mentioned that the price was negotiable. On the questionnaire, subjects were asked how much they would be willing to pay for the second apartment, and this was encoded as Z. The task measure used was (Z - X) / (Y - X), which varies from 0% if the user did not budge from their original requested price, to 100% if they offered the full asking price.

Given literature on the relationship between user personality and preference for computer behavior, we believed subjects might respond differentially to social dialogue based on predisposition. Thus, we also included a composite measure for introversion/extroversion on the questionnaire (PERSONALITY) as in [20].

Extrovertedness was an index composed of seven Wiggins [31] extrovert adjective items: Cheerful, Enthusiastic, Extroverted, Jovial, Outgoing, and Perky.

Introvertedness was an index composed of seven Wiggins [31] introvert adjective items: Bashful, Introverted, Inward, Shy, Undemonstrative, Unrevealing, and Unsparkling.

Results

Full factorial single measure ANOVAs were run, with CONDITION and PERSONALITY as independent variables.

There were no main effects for TRUST, however there was a significant interaction between PERSONALITY and TRUST ($F(1,44)=5.0$; $p<.05$) (see Figure 3). These results indicate that small talk had essentially no effect on how introverts assessed trust but a significant effect on the trust assessment of extroverts; in fact social dialogue seemed to

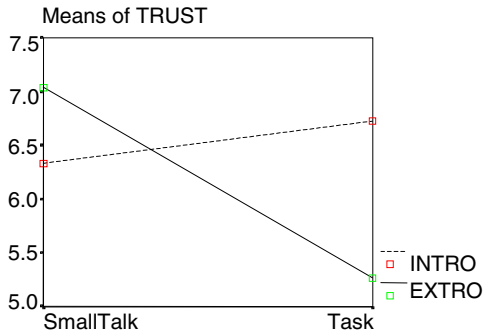

Figure 3: Trust Estimation by introverts & extroverts

be a pre-requisite for establishing the same level of trust for extroverts as that experienced by introverts.

A similar pattern of significant interaction was found between PERSONALITY and several other measures. Extroverts said they felt that REA knew them and their needs better in the SMALLTALK condition, while introverts said that REA knew them better in the TASK condition ($F(1,44)=4.4$; $p<0.05$). Extroverts also said they felt that they knew REA better in the SMALLTALK condition, while introverts said that they knew REA better in the TASK condition ($F(1,44)=5.3$; $p<0.05$). Extroverts also felt the interaction was more natural ($F(1,44)=4.0$; $p<0.06$), satisfying ($F(1,44)=9.6$; $p<0.005$) and successful ($F(1,44)=5.4$; $p<0.05$) with small talk, while introverts said the same of the TASK condition. Finally, extroverts said that REA was more credible in the SMALLTALK condition, while introverts felt she was more credible in the TASK condition ($F(1,44)=3.4$; $p<0.08$).

There was one main effect on CONDITION. Users felt that REA was more engaging in the SMALLTALK condition ($F(1,44)=4.0$; $p<0.06$). There were two main effects on PERSONALITY: extroverts tended to offer more money ($F(1,44)=3.8$; $p<0.07$) and found the interaction more interesting ($F(1,44)=5.3$; $p<0.05$).

No significant effects were found on Amount Willing to Pay for CONDITION. Although we had assumed that there would be a correlation between trust in Rea and this measure, there may be other factors involved in the pricing decision, and we plan to investigate these in the future.

Observation of the videotaped data made it clear that some subjects took the initiative in the conversation, while others allowed REA to lead. Unfortunately, REA is not yet able to deal with user-initiated talk, and so user initiative often led to REA interrupting the speaker. To assess the effect of this phenomenon, we divided subjects into *passive* (below the mean on number of user-initiated utterances) and *initiaters* (above the mean on number of user-initiated utterances) (INITIATIVE). To our surprise, this measure turned out to be independent of intro/extroversion, and to not be predicted by these latter variables (Pearson r = 0.053). Full factorial ANOVAs were again performed on all measures, with CONDITION and INITIATIVE as dependent variables. There were significant interactions between INITIATIVE and several measures. Active users felt that the interaction was more interesting ($F(1,28)=5.2$; $p<0.05$), that REA came to know them better ($F(1,28)=4.4$; $p<0.05$), that they knew REA better ($F(1,28)=14.3$; $p<0.001$) (see Figure 4), and that REA was more of an expert ($F(1,28)=3.5$;$p<0.08$) when she used small talk.

Discussion and Conclusion

Overall we found that users who reach out more towards other people are more susceptible to relationship building and need some relational conversational strategies in order to trust the interface.

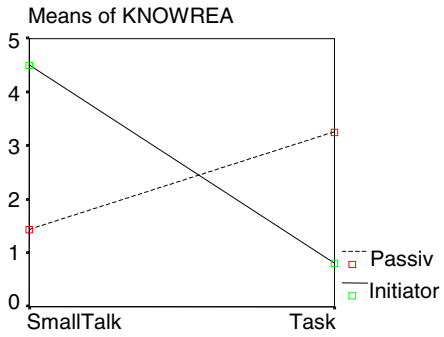

Figure 4: . How well users felt they knew REA by initiaters vs. passive speakers

Relational intelligence includes knowledge of when and how to use language to achieve social goals. This knowledge is crucial for our computational agents if they are to be as effective as people, and if we want people to be able to use our agents easily, efficiently, and cooperatively.

As embodied conversational agents become ubiquitous, the ability for them to establish and maintain social relationships with us will become increasingly important.

We are currently investigating the implementation of other forms of social dialogue and additional relational strategies, as well as expanding the dyadic relationship model used in our discourse planner.

For the moment, however, we have shown that models of social dialogue can be formalized, and that their evaluation demonstrates the importance of the phenomenon to a well-defined subset of users. The study of human-computer relationships is a new field which exists at the nexus of research into human-computer interaction, human social psychology, sociology, and linguistics. The study of how to constitute relationships through language will inform our growing ability to emulate aspects of humans in the service of efficient interaction between humans and machines.

REFERENCES

1. Andre, E., Muller, J., and Rist, T., The PPP Persona: A Multipurpose Animated Presentation Agent, in *Proceedings of Advanced Visual Interfaces*, 1996.
2. Berscheid, E. and Reis, H., Attraction and Close Relationships. *The Handbook of Social Psychology*, D. Gilbert, S. Fiske, and G. Lindzey, Eds. McGraw-Hill, New York , 1998, 193-281.
3. Beskow, J. and McGlashan, S., Olga: a converational agent with gestures, in *Proceedings of IJCAI '97*, 1997.
4. Brown, P. and Levinson, S., Universals in language usage: Politeness phenomena. *Questions and Politeness: Strategies in Social Interaction*, E. Goody, Ed. Cambridge University Press, Cambridge , 1978.
5. Brown, R. and Gilman, A., The pronouns of power and solidarity. *Language and Social Context*, P. Giglioli, Ed. Penguin, Harmondsworth , 1972, 252-282.

6. Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjalmsson, H., and Yan, H., Embodiment in Conversational Interfaces: Rea, in *Proceedings of CHI '99*, (Pittsburgh, PA, 1999), 520-527.
7. Cassell, J., Bickmore, T., Campbell, L., Vilhjalmsson, H., and Yan, H., Human Conversation as a System Framework: Designing Embodied Conversational Agents. *Embodied Conversational Agents*, J. Cassell, J. Sullivan, S. Prevost, and E. Churchill, Eds. MIT Press, Cambridge, MA , 2000.
8. Cassell, J., Sullivan, J., Prevost, S., and Churchill, E., *Embodied Conversational Agents*. MIT Press, Cambridge, 2000.
9. Cheepen, C., *The Predictability of Informal Conversation*. Pinter, New York, 1988.
10. Clark, H. H., *Using Language*. Cambridge University Press, Cambridge, 1996.
11. Dehn, D. M. and Mulken, S. v., *The Impact of Animated Interface Agents: A Review of Empirical Research*, University of Saarland, Saarbrucken, Germany 1999.
12. Fogg, B. J. and Tseng, H., The Elements of Computer Credibility, in *Proceedings of CHI '99*, 1999, 80-87.
13. Goffman, I., On face-work. *Interaction Ritual: Essays on Face-to-Face Behavior*. Pantheon, New York , 1967, 5-46.
14. Jaworski, A. and Coupland, N., *The Discourse Reader*. Routledge, London, 1999.
15. Lester, J., Stone, B., and Stelling, G., Lifelike Pedagogical agents for Mixed-Initiative Problem Solving in Constuctivist Learning Environments, *User Modeling and User-Adapted Interaction 9*, 1-2, 1999, 1-44.
16. Lester, J. C., Voerman, J. L., Towns, S. G., and Callaway, C. B., Cosmo: A Life-like Animated Pedagogical Agent with Deictic Believability, in *Proceedings of IJCAI '97*, 1997.
17. Maes, P., How to do the right thing, *Connection Science Journal 1*, 3, 1989,
18. Malinowski, B., The problem of meaning in primitive languages. *The Meaning of Meaning*, C. K. Ogden and I. A. Richards, Eds. Routledge & Kegan Paul,1923.
19. Moon, Y., *Intimate self-disclosure exhanges: Using computers to build reciprocal relationships with consumers*, Harvard Business School, Cambridge, MA Working paper 99-059, 1998.
20. Nass, C. and Lee, K., Does Computer-Generated Speech Manifest Personality? An Experimental Test of Similarity-Attraction, in *Proceedings of CHI 2000*, (The Hague, 2000), 329-336.
21. Reeves, B. and Nass, C., *The Media Equation: how people treat computers, televisions and new media like real people and places*. Cambridge University Press, Cambridge, 1996.
22. Resnick, P. V. and Lammers, H. B., The Influence of Self-esteem on Cognitive Responses to Machine-Like Versus Human-Like Computer Feedback, *The Journal of Social Psychology 125*, 6, 1985, 761-769.
23. Rickel, J. and Johnson., W. L., Animated agents for procedural training in virtual reality: Perception, cognition, and motor control., *Applied Artificial Intelligence*, 1998.
24. Rickenberg, R. and Reeves, B., The Effects of Animated Characters on Anxiety, Task Performance, and Evaluations of User Interfaces, in *Proceedings of CHI 2000*, (The Hague, 2000), 49-56.
25. Schneider, K. P., *Small Talk: Analysing Phatic Discourse*. Hitzeroth, Marburg, 1988.
26. Spencer-Oatey, H., Reconsidering power and distance, *Journal of Pragmatics 26*, 1996, 1-24.
27. Stone, M. and Doran, C., Sentence Planning as Description Using Tree-Adjoining Grammar, in *Proceedings of ACL*, 1997, 198--205.
28. Svennevig, J., *Getting Acquainted in Conversation*. John Benjamins, Philadephia, 1999.
29. Thorisson, K. R., Gandalf: An Embodied Humanoid Capable of Real-Time Multimodal Dialogue with People, in *Proceedings of Autonomous Agents '97*, 1997).
30. Wheeless, L. and Grotz, J., The Measurement of Trust and Its Relationship to Self-Disclosure, *Human Communication Research 3*, 3, 1977, 250-257.
31. Wiggins, J., A psychological taxonomy of trait-descriptive terms, *Journal of Personality and Social Psychology 37*, 3, 1979, 395-412.