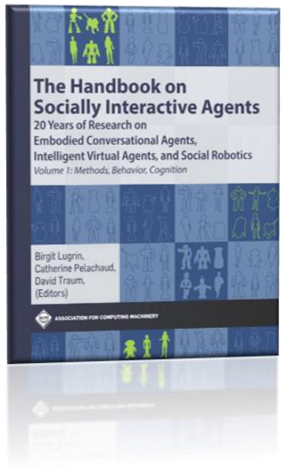
Rapport Between Humans and Socially Interactive Agents

Jonathan Gratch and Gale Lucas



Author note:

This is a preprint. The final article is published in "The Handbook on Socially Interactive Agents" by ACM books.

Citation information:

Gratch, J., and Lucas, G. (2021). Rapport Between Humans and Socially Interactive Agents. In B. Lugrin, C. Pelachaud, D. Traum (Eds.), *Handbook on Socially Interactive Agents – 20 Years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics*, Volume 1: Methods, Behavior, Cognition (pp. 433-462). ACM.

DOI of the final chapter: 10.1145/3477322.3477335

DOI of volume 1 of the handbook: 10.1145/3477322

Correspondence concerning this chapter should be addressed to Jon Gratch (gratch@ict.usc.edu) or Gale Lucas (lucas@ict.usc.edu)

Rapport Between Humans and Socially Interactive Agents

Jonathan Gratch, Gale Lucas

Institute for Creative Technologies, University of Southern California

12014 Waterfront Dr., Playa Vista, CA 90094

gratch@ict.usc.edu; lucas@ict.usc.edu

1. Introduction

Think back on a time when you saw speakers engrossed in conversation. Even without hearing their words you can sense their connection. The participants seem tightly enmeshed in something like a dance. They rapidly resonate to each other's gestures, facial expressions, gaze and patterns of speech. Such behavior has been studied by many names including social resonance (Duncan et al., 2007; Kopp, 2010), interpersonal adaptation (Burgoon, Stern, & Dillman, 1995), entrainment (Levitan, Gravano, & Hirschberg, 2011), interactional synchrony (Bernieri & Rosenthal, 1991), social glue (Lakin, Jefferis, Cheng, & Chartrand, 2003), immediacy behaviors (Julien, Brault, Chartrand, & Bégin, 2000), and positivity resonance (Fredrickson, 2016). In this chapter, we follow the terminology of Tickle-Degnen and Rosenthal (1990) and refer to this seemingly automatic attunement of positive emotional displays, gaze, and gestures as *rapport.*

Regardless of the difference in focus or definition, these bodies of research agree that, once established, rapport influences a wide range of interpersonal processes. It plays a crucial role in the establishment of social bonds (Tickle-Degnen & Rosenthal, 1990), promoting or diffusing conflict (Lanzetta & Englis, 1989), persuasion (Bailenson & Yee, 2005; Fuchs, 1987) and in the establishment of identity (Mead, 1934). Perhaps more importantly, rapport leads to beneficial outcomes across a wide range of practical interpersonal problems. It fosters success in negotiations (Drolet & Morris, 2000; Goldberg, 2005), improves workplace cohesion (Cogger, 1982), enhances psychotherapeutic effectiveness (Tsui & Schultz, 1985), elevates test performance in classrooms (Fuchs, 1987) and raises the quality of child care (Burns, 1984). To paraphrase Tickle-Degnen and Rosenthal (1990), this chapter explores how intelligent virtual clinicians can develop rapport with patients, robot sales personnel can use it to make a deal, and socially interactive agents can try to predict from it the future of their relationship with a user.

In people, such unfolding and resonating patterns of behavior arise without conscious deliberation (Bargh, Chen, & Burrows, 1996; Lakin et al., 2003) but they are hardly fixed reflexes. The same individual that smiles reflexively to the smile of a friend may frown to the smiles of his opponent (Lanzetta & Englis, 1989). The way such interpersonal patterns unfold depends on a host of intraand interpersonal factors including the relative power of individuals in an interaction – with weaker partners tending to mimic the more powerful but not vice versa (Tiedens & Fragale, 2003); prior expectations – with unexpectedly positive behaviors producing more favorable outcomes than expected ones (Burgoon, 1983); and conformity to social norms such as reciprocity (Roloff & Campion, 1987) or the appropriate use of expressing negative emotions in a particular setting (Adam, Shirako, & Maddux, 2010). More generally, these patterns both arise from and help to redefine an evolving affective relationship between individuals (Parkinson, 2013).

Building "rapport agents" that can engage people in this unfolding emotional dance is a fascinating prospect with potentially profound practical and scientific consequences. In terms of applications, socially interactive agents can use rapport to enhance disclosure in mental health screenings (Lucas, Gratch, King, & Morency, 2014) and robots can use it to improve learning in primary school students (Lubold, Walker, Pon-Barry, & Ogan, 2018). In terms of science, computational models of rapport can enhance social theory by allowing systematic manipulation of nonverbal patterns in ways that are difficult or impossible for human confederates in social science experiments. Such experimental control enables studies that can definitively establish if such patterns *cause* these effects or merely reflect them (see Bailenson, Beall, Loomis, Blascovich, & Turk, 2004; Bente, Kraemer, Petersen, & de Ruiter, 2001; de Melo, Carnevale, & Gratch, 2014; Forgas, 2007; Gratch et al., 2006; Hoegen, Schalk, Lucas, & Gratch, 2018). Indeed, this chapter reviews findings that synthesized patterns of nonverbal behavior cause predicted changes in the impressions and behaviors of human subjects in laboratory studies. These studies give insight into the factors that promote these effects, and show how these insights can translate into computer-based applications to health, commerce and entertainment.

We first review rapport theory and why it is important for human-machine systems. We next highlight some of the technical and methodological advances that have allowed the creation of machines that establish rapport with people. Many groups have explored these capabilities in both virtual and robotic agents and this chapter will not give full justice to this vibrant body of research. Rather, we will focus on our own modest contributions to this field. Finally, we will review empirical findings that human-machine rapport has important benefits, that computational models can inform psychological theories of rapport, and that rapport supports important societal outcomes, such as addressing human mental health.

2. Rapport Theory

Rapport is studied across a range of scientific disciplines and domains for its role in fostering emotional bonds and prosocial behavior. This includes research on teaching (Bernieri, 1988), psychotherapy (Charny, 1966), sales (Gremler & Gwinner, 2000) and interrogations (Abbe & Brandon, 2013), to name but a few. Research on rapport is also diverse in the behaviors and modalities addressed. Most research, and the focus of this chapter, emphasizes nonverbal behaviors such as smiles and postures. But other work explores how feelings of rapport arise from, and shape the content of conversations (Friedberg, Litman, & Paletz, 2012), conversational mechanisms such as turn-taking (Cassell, Gill, & Tepper, 2007), or even the convergence of physiological processes such as respiration (McFarland, 2001) and pupil dilation (Kang & Wheatley, 2017).

In their seminal article, Tickle-Degnen and Rosenthal (1990) equate rapport with subjective experience (e.g., individual report that they “clicked” or have a sense of “chemistry”), but also with three essential elements of nonverbal behavior. First are signals of *positivity* such as smiles, nods of agreement, forward trunk lean, uncrossed arms and open body posture. These behaviors indicate mutual liking and approval. The second nonverbal element is signals of *mutual attentiveness* such as shared gaze, direct body orientation and “listening behaviors” such as backchannels (verbal “uh-huh” or quick nods) that convey that participants are actively attending and understanding each other. The third essential element is *coordination*. This refers to dyadic/group behaviors that convey parties are functioning as a coordinated unit, such as postural mimicry or interactional synchrony. Tickle-Degnen and Rosenthal further emphasize that feelings and behavior are tightly linked. They argue for strong consistency between self-reported rapport from each participant in a conversation, the presence of nonverbal components, and third-party judgments. In this sense, rapport is easy to assess -- observers can reliably predict if people will report feelings of rapport by watching their nonverbal behavior (Ambady, Bernieri, & Richeson, 2000) – and these judgments and feelings predict positive interpersonal outcomes.

Rapport also has a dynamic aspect. In the short term, as people begin to establish rapport in a conversation, there is a tendency for behaviors to become more coordinated over time through a process sometimes called *entrainment* (e.g., Borrie, Barrett, Willi, & Berisha, 2019). This can include increased postural and gestural mimicry (Bergmann & Kopp, 2012), convergence of acoustic and prosodic features (Levitan & Hirschberg, 2011), and shared syntax and word use (Brennan, 1996). Over longer time periods (i.e., across multiple conversations), Tickle-Degnen and Rosenthal posit that the nonverbal elements of rapport shift in importance as the relationship between interlocutors deepens and strengthens (see Figure 1). In particular, the need to signal positivity becomes less important as mutual-liking becomes assumed. Although not reflected in Tickle-Degnen and Rosenthal’s model, other research suggests that the appearance of coordination may diminish (at least from the perspective of classical turn-taking models). For example, as participants become more comfortable with each other, barging in and overlapping speech becomes far more common (Coates, 1989).

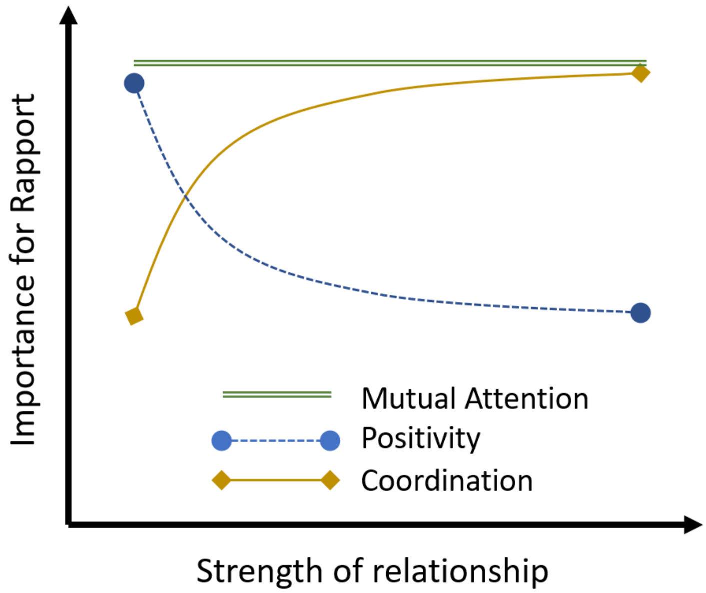


Figure 1: The importance of elements of rapport as relationships deepen (based on Tickle-Degnen & Rosenthal, 1990)

Most psychological research on rapport examines contexts where both parties are actively speaking, but rapport can be established even during a monolog. During monologues, listeners can convey their positivity, mutual attention and coordination through *backchannels* (Yngve, 1970). In linguistics, these are behaviors such as verbal interjections (e.g., "yeah", "right", "uh-huh") or nonverbal signals such as nods, smiles and other facial expressions such as eyebrow raises or expressions of sympathy. These behaviors can increase speaker comfort, fluidity and even change the nature of the monolog. Indeed, Jane Bavelas has emphasized that, through these behaviors, listeners become co-narrators (Bavelas, Coates, & Johnson, 2000).

When it comes to computational models, most systems have emphasized "early rapport" (as shown in Figure 1) and active-listening behaviors such as backchannels. This is probably due to limitations in the current state-of-the-art in conversational agents. Although rapidly advancing, dialog systems are still limited, especially with regard to the sort of continuous and predictive parsing of speech needed to match human conversational fluidity (DeVault, Sagae, & Traum, 2011). Thus, many systems focus on the illusion of understanding by responding to more easy to recognize surface features of speech (Ward & Tsukahara, 2000) and rarely entrain their behavior to participants across multiple interactions. For this reason, we will focus the remainder of this chapter on systems that emphasize active listening approaches to establishing rapport.

3. History and Overview of Rapport Agents

The practical benefit of rapport between people has spurred a wide array of research into techniques that can reproduce this interpersonal state within human-machine interactions. We trace the evolution of a fifteen-year effort within our laboratory to endow machines with the elements of rapport, beginning with the “Rapport Agent” (that engaged in simple listening behaviors), and ending with SimSensei, an intelligent virtual agent that integrates multi-modal rapport behaviors with natural language dialog. During this period, the technology underlying such rapport agents advanced from hand-crafted rules (based on theoretical observations) to machine learning approaches and we innovated new ways to collect and annotate the data that machine learning approaches demand, such as Parasocial Consensus Sampling. In terms of applications, this project began as a theoretical exercise to study if human-machine rapport was even possible. As these methods have matured, they are increasingly applied to practical applications, particularly in health interviews (Lucas et al., 2014) and behavior change (see Chapter 25).

A rule-based approach

The Rapport Agent (Gratch et al., 2006), illustrated in Figure 2, was designed to evoke subjective feelings and behavioral consequence of rapport with human participants in a restricted set of social interactions we refer to as *quasi-monologs.* In a quasi-monolog, the human does most of the talking and the agent primarily prompts human speech and provides attentive listening feedback. The Rapport Agent was originally developed to replicate findings by Duncan and colleagues on listening behaviors (Welji & Duncan, 2004). In their studies, a speaker (the narrator) retells some previously observed series of events (e.g., the events in a recently-watched video) to some listener. In our case, the listener is a computer-generated character that has been programmed to produce (or fail to produce) the type of nonverbal feedback seen in rapportful conversations.

The agent was crafted by three visiting interns, Francois Lamothe and Mathieu Morales from the French Military Academy and Rick van der Werf from Twente University. The implementation focused on the observation that, when it comes to active listening, feelings of rapport correlate with simple contingent listening behaviors such as the use of backchannels (nods, elicited by speaker prosodic cues, that signify the communication is working), postural mirroring, and mimicry of certain head gestures, e.g., gaze shifts and head nods (Chartrand & Bargh, 1999; Ward & Tsukahara, 2000; Yngve, 1970). To identify backchannel opportunity points, Lamothe and Morales created an acoustic tool called Luan, building on techniques proposed by Nigel Ward (Ward & Tsukahara, 2000). To identify nonverbal behaviors, van der Werf utilized the Watson software package developed by Louis-Philippe Morency (Morency, Sidner, Lee, & Darrell, 2005). This component uses a seated participant’s head position to detect posture shifts and use head orientation to recognize head nods and shakes.

Figure 2: The first-author's daughter interacting with the Rapport Agent.

Figure 3 illustrates the Rapport Agent architecture. Participants told stories to a virtual agent rendered in a game engine and animated with the SmartBody animation system (Kallmann & Marsella, 2005; Thiebaux, Marshall, Marsella, & Kallmann, 2008). Features recognized from speech and vision were passed to a simple authorable rule-based system which communicated messages in the Behavior Markup Language (Kopp et al., 2006) to SmartBody. The original character, illustrated in Figure 2, was intended to look like Brad Pitt, an obvious failure, but this was the origin of a long line of characters named Brad emanating from our laboratory.

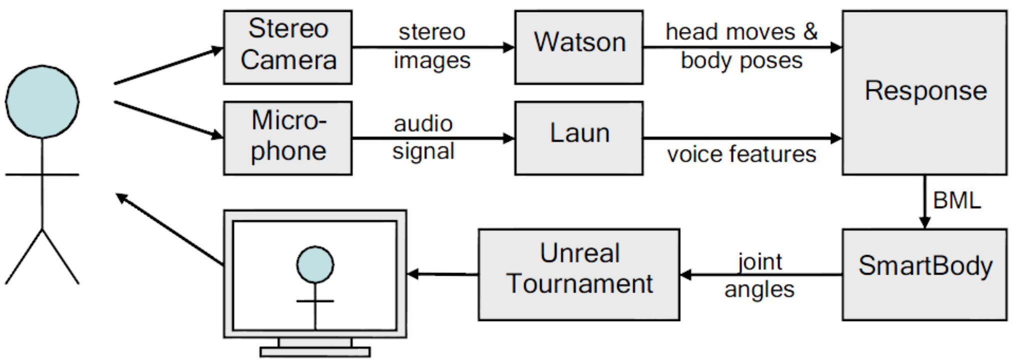


Figure 3: Rapport Agent architecture.

Though simple and based on hand-crafted rules, the Rapport Agent proved remarkably effective. Other groups reported success with alternative rule-based systems for virtual (Truong, Poppe, & Heylen, 2010) or robotic agents (Fujie, Ejiri, Nakajima, Matsusaka, & Kobayashi, 2004). As discussed in the next section, such agents create subjective feelings of rapport and several beneficial changes in participant behavior (fostering more fluent speech and eliciting longer and more intimate stories). Such systems also served as an important empirical tool, helping to establish the importance of all three theoretically-positive nonverbal elements of rapport.

Machine Learning

The Rapport Agent, and similar agents at the time, relied on hand-crafted rules due to the lack of good annotated datasets on rapport. One consequence of the project was the creation of a reasonably large corpus (by standards of the time) of human to human conversations that were rated by both parties for perceptions of rapport and annotated (both manually and automatically) for nonverbal behaviors associated with rapport (this data is publicly available at https://rapport.ict.usc.edu/). This data afforded the opportunity to explore machine learning approaches to replace the hand-crafted rules.

Our initial machine learning efforts focused on predicting backchannel opportunity points. Figure 4 illustrates our approach that was spearheaded by Louis-Phillipe Morency and another University of Twente intern, Iwan de Kok (Morency, de Kok, & Gratch, 2008, 2010). The approach encoded multimodal features from a speaker using an encoding dictionary to represent different possible temporal relationships between speaker behaviors and listener responses. For example, a step function hypothesizes that the backchannel is likely whenever the speaker's feature is present, whereas a ramp function hypothesizes that the backchannel is most likely when the speaker's feature is *first* present. These features were then passed to a supervised machine learning approach based on conditional random fields (Lafferty, McCallum, & Pereira, 2001).

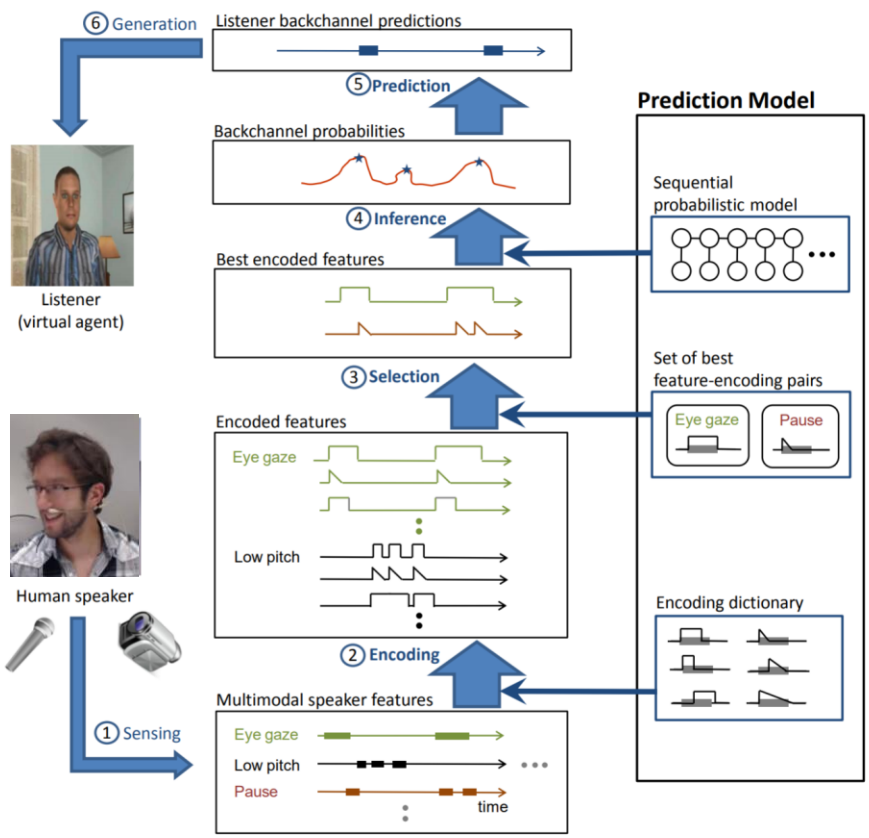


Figure 4: Architecture for multimodal backchannel prediction. The system (1) senses audiovisual information from a speaker. These features are (2) encoded using a dictionary that represented alternative possible temporal relationships between speaker behaviors and listener responses. After feature selection (3), these were passed to a machine learning classifier (4) which generates probabilistic predictions (5).

The learning approach yielded considerable improvement over our hand-crafted rules. An analysis of the learned model emphasized the importance of multi-modality: back channels were elicited by prosodic features, as predicted by Ward and Tsukahara (2000), but also depended on visual features (listeners would tend to backchannel when the speaker paused and looked at them). Thus, at least in terms of backchannels, this research showed that elements of rapport could be automatically learned from data.

Parasocial Consensus Sampling

If the same storyteller tells the same story to different audiences, they may get quite different reactions. Some audiences may be engaged and show active listening, which in turn will feed back and help the speaker tell a better story (Bavelas et al., 2000). Others may sit impassively. Even when engaged, there may be considerable variability in how feedback is offered. As learning models began to capture the basics of rapport, an emerging question became how to explain this variability.

Lixing Huang, a PhD student in our laboratory, proposed an interesting approach to examine and analyze variability in listening behavior. Rather than focusing on dyads, he asked whether it was possible to get the responses of many listeners to a single speaker. Around the same time, de Kok and Heylen were raising the same question (de Kok & Heylen, 2011). But Dr. Huang further wondered if people really needed to be in a real conversation to provide useful data about how they would respond. In communication theory, Horton and Wohl (Horton & Wohl, 1956) introduced the term *parasocial interaction* to refer to the observation that people often respond to media, like cinema or television, as if they are really engaged in a social interaction with them (think of the time you yelled at a TV pundit you disagreed with). Building on this idea, he created a framework where crowdworkers could watch a storyteller and "act out" the types of responses they would show if they were in a real social interaction. The technique turned out to be surprisingly effective and allowed the inexpensive creation of large corpora of rapport behavior (Huang, 2013).

Huang termed his method *parasocial consensus sampling (PCS).* The idea was to have multiple coders signal some form of feedback (he began with simple backchannels but then moved to turn-taking and emotional feedback). This would elicit an entire distribution of responses to each speaker. As seen in Figure 5, coders highlight multiple opportunities to provide feedback but only reached consensus on the third opportunity. The consensus response could be used to train a classifier (Huang, Morency, & Gratch, 2010). However, the variability was interesting in its own right. One could use the amount of consensus as a measure of how salient certain features were in eliciting feedback. For example, the word "bothering" seemed to trigger a stronger response. Further, one can examine variability across different "types" of individuals. For example, extraverts are more likely to provide feedback during opportunity points than introverts (Huang & Gratch, 2013).

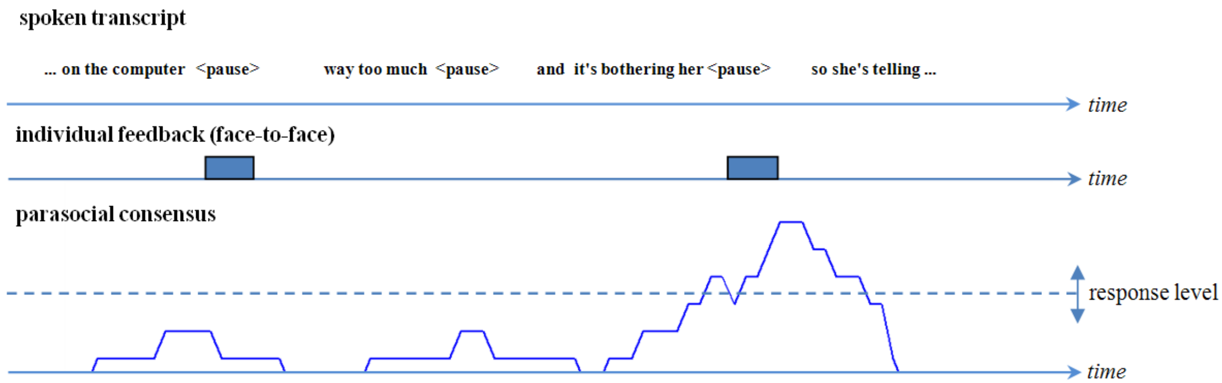


Figure 5: Illustration of the parasocial consensus sample of a particular storyteller. The first line shows the speakers transcript. The second line shows the backchannels provided by the original "real" listener. The bottom line shows the distribution of responses by PCS coders.

The PCS technique was applied to backchannels but also turn-taking and emotional feedback. The resulting learned models yielded insight into these conversational processes. For example, people allow the speaker to keep the turn longer when they pause if the speaker looks away. In terms of emotional feedback, mimicry was far more likely to occur during backchannel opportunity points. The resulting models were incorporated into an improved version of the Rapport Agent that could interview people, with a fixed script of questions, pause the appropriate length of time before moving on, and provide the semblance of active and empathetic feedback (Huang, Morency, & Gratch, 2011).

SimSensei

Perhaps the biggest limitation (and concern) about the above-mentioned rapport agents is that they can create the illusion of understanding without any actual understanding. While this can be a useful experimental tool for studying rapport, it can be problematic in an actual application. Our most recent research on the SimSense agent attempted to correct this by integrating the capabilities of the Rapport Agent with a state-of-the-art dialog system (DeVault et al., 2014).

SimSensei was designed to create interactional situations favorable to the automatic assessment of distress indicators, defined as verbal and nonverbal behaviors correlated with depression, anxiety or post-traumatic stress disorder (PTSD). The agent would simulate a mental health screening by interviewing participants about the challenges they may be facing in life. The agent implemented dialog-management using the FLoReS dialog manager (Morbini et al., 2014). This grouped interview questions into several phases including initial chit-chat designed to enhance rapport, followed by a set of diagnostic questions and ending with a "cool down" phase. Active dialog management allowed the agent to ask follow-up questions. It also allowed the agent to incorporate verbal and valanced backchannel feedback. For example, rather than simply nodding at backchannel opportunity points, the agent could respond "that's great" or "I'm sorry" depending on what it recognized. In practice, this feedback had to be used sparingly due to the current limits of natural language understanding. Nonverbal feedback and gestures were substantially improved over the above-mentioned systems by using the Cerebella nonverbal behavior generation system (Marsella et al., 2013). As discussed in our empirical findings, incorporating rapport behaviors into a computer program proved an effective way to reduce participants' fear of being judged and increased comfort in disclosing symptoms of mental illness.

4. Empirical Findings

The above-described rapport agents exhibit the nonverbal behaviors that occurs in natural rapportful conversation, but are people influenced by computer-generated behaviors? In a series of empirical studies, we have used rapport agents as methodological tools to explore the role of nonverbal behaviors in human and human-computer interaction. These studies clearly demonstrate that synthetic rapport behaviors alter participant feelings, impressions and behavior, and that the strength of such effects are mediated by the same elements of positivity, contingency and mutual attention that Tickle-Degnen and Rosenthal have posited for face-to-face interactions between people. Further, as posed by Nass and Reeves in their *media equation* (Reeves & Nass, 1996), such effects seem to occur even when participants know that such behaviors are "merely" produced by a computer.

In the remainder of this chapter, we outline the basic experimental paradigm by which we explore such questions and then summarize the findings of several studies. We highlight different aspects of these findings. First, we describe the consequences, subjective and behavioral, that such contingent nonverbal behaviors have on human participants. Next, we discuss some of the factors that seem to mediate these effects. These include properties of the agent's behavior, dispositional factors on the part of participants, and situational factors surrounding the interaction such as if participants believe they are interacting with another participant or a computer program.

Experimental Paradigm

Figure 2 illustrates a typical experimental setup. Participants would sit in front of a rapport agent and be prompted (either by the experimenter or by the agent itself) to retell some previously experienced situation – in one series of experiments, participants watched a short video and were then instructed to retell it to the agent in as much detail as possible. More recent experiments, for example with the SimSensei system, involved more interactive dialog interviewing participants about their real-life experiences. In either situation, the agent displays some form of nonverbal feedback while the participant speaks, with the exact nature of the feedback being dictated by the specific experimental manipulation.

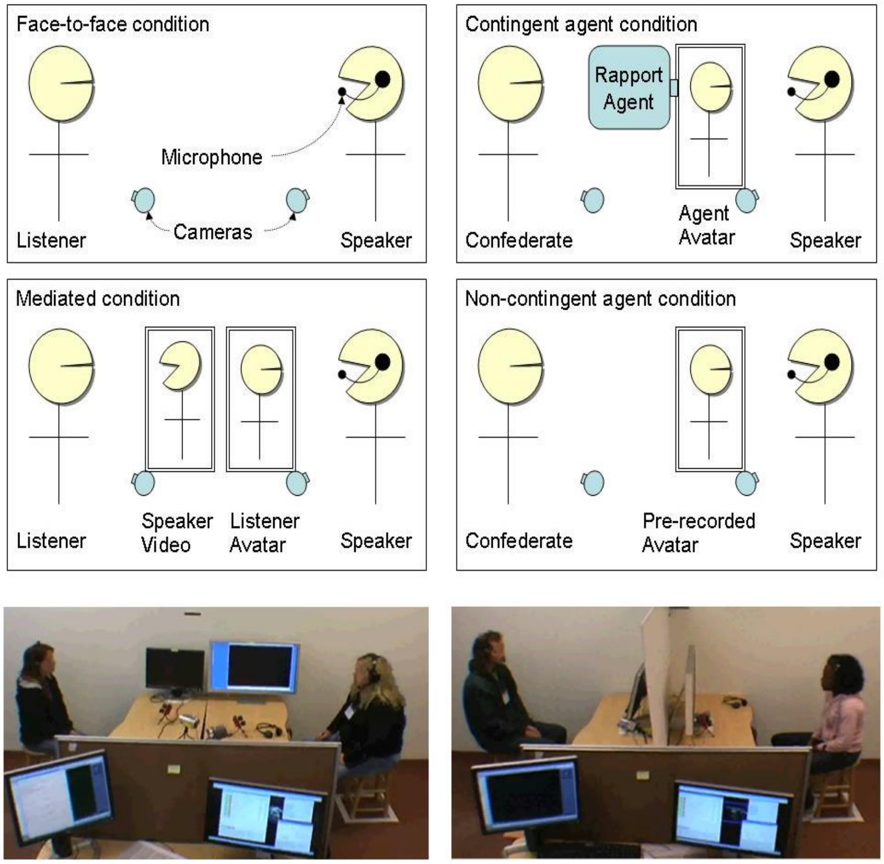


Figure 6. Graphical depiction of the four conditions. The actual face-to-face condition is illustrated on the lower left and the setup for the other three conditions on the lower right.

In these studies, participant rapport was assessed through a variety of subjective and behavioral measures. Subjective measures included scales assessing rapport, social presence (Biocca & Harms, 2002), helpfulness, distraction and naturalness.[1] Behavioral measures included the length people speak (as a measure of engagement), the fluency of their speech (e.g., how many repeated words, broken words, and filled pauses per minute), the depth of their disclosure, as well as facial expressions and eye-gaze patterns. Behavioral measures were assessed through a mixture of automatic annotation techniques and hand annotations by multiple coders.

[1] For the evolution of our own rapport scales, see https://rapport.ict.usc.edu/data/rapport-2006/RapportScales_summary_sKang_updated.pdf

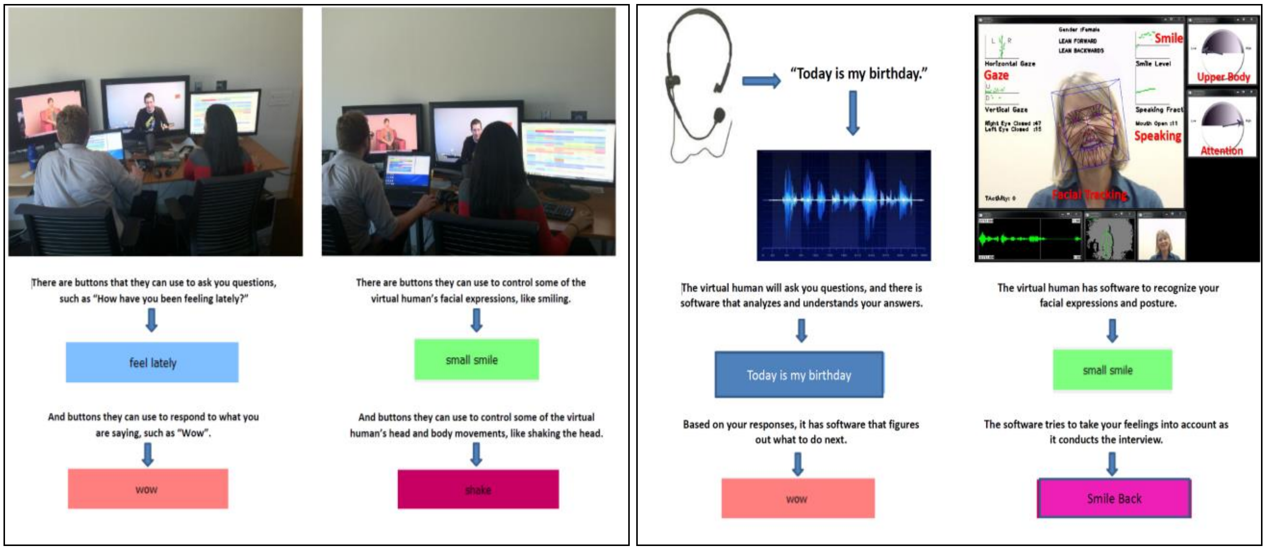


Figure 7. Graphical depiction of the two belief conditions. Participants were told that they were interacting with a virtual agent controlled by human operators (left) or told it was driven by a computer program that used speech and vision (right).

Part of the power of computational models is we can systematically manipulate aspects of the agent's appearance and nonverbal behavior, as well as prior beliefs about the situation. For example, Figure 6 (from Gratch, Wang, Gerten, & Fast, 2007) illustrates a study that examined the impact of appearance (human vs. computer-generated human), behavior (human-generated vs. computer-generated) and contingency ("properly" timed vs. random listener feedback). Going clockwise in this figure from the upper-left, this experiment compared face-to-face interaction (in which a visible human participant displayed natural listening behaviors), the Rapport Agent (which exhibited computer-generated behavior and appearance), a non-contingent version of the Rapport Agent (which exhibited behaviors identical to the Rapport Agent in terms of their frequency and dynamics, but not contingent on the behaviors of the speaker), and a "mediated agent" (in which a real participant's listening behaviors were displayed on a computer-generated avatar). Other research has manipulated beliefs about the situation such as whether the participant believes they are interacting with another human (represented by an avatar) or if they are interacting with an autonomous agent (e.g., Lucas et al., 2014; von der Pütten, Krämer, Gratch, & Kang, 2009). These beliefs were shaped by prior instruction. For example, in an experiment involving SimSensei, participants were provided a detailed explanation of how the avatar or agent software worked, as seen in Figure 7.

Social effects of the rapport agent

The picture emerging from a series of studies is that the Rapport Agent elicits beneficial social effects from human participants similar to what can be found in rapportful face-to-face interactions. Broadly, a rapport agent's behavior shapes subjective feelings:

- Greater feelings of self-efficacy (Kang, Gratch, Wang, & Watt, 2008)
- Less tension (Wang & Gratch, 2010) and less embarrassment (Kang, Gratch, Wang, & Watt, 2008)
- Greater feelings of rapport (Wang & Gratch, 2010)

- A greater sense of mutual awareness (von der Pütten, Krämer, & Gratch, 2009)
- Greater feelings of trustworthiness on the part of the agent (Kang, Gratch, Wang, & Watt, 2008)

And a rapport agent's behavior shapes human behavior:

- More disclosure of information including longer interaction times and more words elicited (Gratch et al., 2006; Gratch et al., 2007; von der Pütten, Krämer, & Gratch, 2009; Wang & Gratch, 2010)
- More fluent speech (Gratch et al., 2006; Gratch et al., 2007; von der Pütten, Krämer, & Gratch, 2009; Wang & Gratch, 2010)
- More mutual gaze (Wang & Gratch, 2010)
- Fewer negative facial expressions (Wang & Gratch, 2009b)
- Improvement in performance when the agent supervises them taking an academic test (Karacora, Dehghani, Krämer, & Gratch, 2012; Krämer et al., 2016)
- Satiates belonging needs among those with chronically high need to belong, thus reducing intentions to seek social interaction with others after interacting with the rapport agent (Krämer, Lucas, Schmitt, & Gratch, 2018)

But besides verifying these general effects, our research has sought to illuminate factors that mediate or moderate these relationships and more generally, to explore the validity of alternative theoretical constructs for interpreting these results and guiding future agent design. We organize this research review around three basic questions. First, what properties of agents are necessary or sufficient for promoting rapport? Second, what characteristics of people lead them to be more or less influenced by the agent's behavior? Finally, we consider the more general question of the usefulness of social psychological theory (which was developed to explain human-to-human interaction) as a framework for guiding the design of computer systems. Blascovich (2014) suggests that interactions might unfold very differently depending on whether people believe they are interacting *with computers* or *through computers* (i.e., the Rapport Agent might have different social effects depending on if participants believed its behavior was generated by a computer or if they believed the behavior corresponded to the movements of an actual human listener). This last question, depending on the answer, could have profound effects for the value of interdisciplinary research on social artifacts.

Characteristics of the agent that impact rapport

If we adopt the former perspective and apply social psychological theory directly, Tickle-Degnen and Rosenthal's theory argues that three broad characteristic of agent behavior should promote the establishment of rapport between participants and agents. These factors include *positivity,* meaning that rapport will be enhanced by positive nonverbal cues including positive facial expressions and encouraging feedback such as head nods; *coordination*, meaning that rapport will increase as the behavior of one participant is perceived as contingent upon (i.e., causally related to) the behavior of the other; and *mutual attentiveness*, meaning that rapport will increase as participants attend to each other nonverbally, for example through mutual gaze. We also consider an additional factor: *anonymity*, the sense that one's identity is protected. While direct face-to-face human interactions does not -as a rule- allow for anonymity (and it was therefore not pertinent for Tickle-Degnen and Rosenthal), research with other means of communication suggest that it should also promote the establishment of rapport between participants and agents. Thus, our empirical studies have sought to manipulate these four factors independently and observe their impact on rapport and participants' subsequent behaviors, e.g., disclosure of personal information.

Positivity

In face-to-face conversations, positivity is conveyed through a variety of nonverbal signals such as facial expressions and head nods. To explore the impact of positivity, we have operationalized this dimension in terms of the presence or absence of head nods and facial expressions.

Our findings illustrate that the presence of listener nods significantly enhances feelings and behavioral manifestations of rapport (Wang & Gratch, 2010). The strength of this effect seems moderated by the perceived contingency of the nods (Kang, Gratch, Wang, & Watt, 2008) and dispositional factors of the participant (von der Pütten, Krämer, & Gratch, 2010) as will be discussed below. We also explored whether participant facial expressions could be indicators of rapport, specifically via communicating positivity (Wang & Gratch, 2009b). We looked at participants' facial expressions which were analyzed using the Facial Action Coding System (FACS) and the Computer Expression Recognition Toolbox (Whitehill, Bartlett, & Movellan, 2008). In both of human-to-human and human-to-agent interactions, more positive facial expressions were associated with greater rapport. Most recently, we have examined how an agent that mimics facial expressions in the context of a prisoner's dilemma game can foster rapport (Hoegen et al., 2018). Collectively, these findings are in line with Tickle-Degnen and Rosenthal's predictions.

Coordination

Coordination occurs when two participants' behaviors are mutually related in a timely and positive manner. We have operationalized this factor by manipulating whether behaviors (such as nods or posture shifts) were generated in response to the participant's behavior (a coordinated condition) or in response to some unrelated factor (an uncoordinated condition). For example, in one study, we created non-contingent behavior by showing a participant the same nonverbal behavior that was generated for the previous participant. This "yoked" experimental design only breaks contingency, and still ensures a similar frequency and distribution of behaviors are generated.

Overall, participants exhibit more subjective and behavior characteristics of rapport when coordination is present. For example, breaking the contingency of nonverbal feedback leads participants to talk less and produced more disfluent speech (Gratch et al., 2007). These effects were especially strong in participants that scored high in a scale of social anxiety, as such participants, in addition to these behavioral effects, feel less subjective rapport and greater embarrassment (Kang, Gratch, Wang, & Watt, 2008). We further explored the effects of coordination on learning and found some evidence that coordination helps improve a speakers retention of the event they are discussing (Wang & Gratch, 2009a). Collectively, these findings are in line with Tickle-Degnen and Rosenthal's predictions.

Mutual attentiveness

Mutual attentiveness occurs in a conversation when both participants attend closely to each other's words or movements. In our empirical research, we have operationalized this concept in terms of gaze – for example, are participants looking at each other during a conversation. Prior research in gaze suggests there should be a curvilinear relationship between gaze and feelings of rapport. In other words, continuously staring at another person or completely avoiding their gaze would tend to be disruptive but, short of these extremes, rapport should be enhanced by more visual attention. Consistent with this, we found a similar curvilinear relationship between gaze and rapport. For example, an agent that continuously gazed at the participant without other accompanying timely positive gestures (e.g., nodding) caused more distractions, less rapport and more speech disfluency in storytelling interaction (Wang & Gratch, 2010). Similarly, an agent that failed to maintain gaze with the participant was equally disruptive. Collectively, these findings are in line with Tickle-Degnen and Rosenthal's predictions.

Anonymity

People feel a sense of anonymity when they believe that their identity is protected. In our empirical research, we have found that manipulating a sense of anonymity can encourage positive downstream consequences of rapport such as disclosure of personal information. Indeed, research has shown that greater feelings of rapport lead people to disclose more (Burgoon, Guerrero, & Manusov, 2016; Dijkstra, 1987; Gratch et al., 2007; Hall, Harrigan, & Rosenthal, 1995; Miller, Berg, & Archer, 1983). Like rapport, anonymity also increases disclosure of personal information (for a review, see Weisband & Kiesler, 1996). Because they are anonymous, computer programs give people a "sense of invulnerability to criticism, an illusion of privacy, the impression that responses 'disappear' into the computer" (Weisband & Kiesler, 1996), and therefore because rapport agents might be perceived as anonymous, they have the potential to increase disclosure of personal information as well (Baker, 1992; Beckenbach, 1995; Joinson, 2001; Sebestik, Zelon, DeWitt, O'Reilly, & McGowan, 1988; Thornberry, Rowe, & Biggar, 1991). Rapport agents therefore have the potential to use both their rapport-building capabilities and the sense of anonymity they foster as two "routes" to encourage people to open up and disclose more personal information.

Indeed, when rapport agents use rapport-building behaviors contingently, they are able to prompt disclosure from interviewees (Gratch, Kang, & Wang, 2013). Even when used as interviewers in clinical settings, virtual agents get users to share more personal information when people believe they are operated by a computer than a human (Lucas et al., 2014; Mell, Lucas, & Gratch, 2017; Pickard, Roster, & Chen, 2016). Because conversation with intelligent virtual agents are typically experienced as more anonymous than similar conversations with humans, users seem to be more comfortable disclosing about highly sensitive topics, like their mental health, and on questions that could lead them to admit something stigmatized or otherwise negative. For example, during a clinical interview with a virtual character, participants disclose more personal details when they are told that the character is controlled by artificial intelligence than when they are told that the character is operated by a person in another room (Lucas et al., 2014). In 2016, Pickard and colleagues reported that individuals are more comfortable disclosing to an automated interviewer than a human interviewer.

However, additional research in this line demonstrates the value of having the anonymous agent build rapport. We also examined (Lucas et al., 2017) whether a rapport agent could increase disclosure of mental health symptoms among active-duty service members and veterans. Replicating prior work showing that service members report more symptoms of PTSD when they anonymously answer the Post-Deployment Health Assessment (PDHA) symptom checklist compared to the official PDHA that goes on their permanent record (Warner et al., 2011), service members reported more symptoms during a conversation with an intelligent virtual interviewer than on the official PDHA. This demonstrated the importance of anonymity for disclosure of personal information. However, across two studies, active duty and retired service members also reported more symptoms to a rapport agent (SimSensei) than on an *anonymized version* of the PDHA. Given that the virtual intelligent interviewer and anonymized version of the PDHA were equally anonymous but only the virtual intelligent interviewer could evoke rapport, this work establishes the idea that rapport has an impact on self-disclosure above and beyond anonymity. Therefore, socially intelligent interviewers that build rapport may provide a superior option to encourage disclosure of personal information. Pragmatically, this finding makes the case for taking advantage of the value that rapport-building holds for increasing disclosure of information rather than just relying on anonymity.

An important qualification to the impact of anonymity may be that it depends on the social situation. Specifically, in contexts where they might be judged, people show more social effects with an agent than an avatar. Indeed, as described above, research has found that people respond more socially to agents than avatars in settings where they feel judged: clinical interviews about their mental health (Lucas et al., 2014; Lucas et al., 2017; Pickard et al., 2016; Slack & Van Cura, 1968), but they also do so when interviewed about their personal financial situation (Mell et al., 2017). Some of this work shows that people are more comfortable with agents *only* if they are concerned about being judged (Pickard et al., 2016).

In contrast, when rapport agents engage with users on other tasks where judgement is less likely, the results are more complicated. First, during persuasive conversations, rapport agents are more persuasive than human-operated agents as long as they give off cues that they are competent (Khooshabeh & Lucas, 2018; Lucas, Lehr, Krämer , & Gratch, 2019). Indeed, if their competence is called into question (e.g., by making repeated errors), rapport agents have less social influence than agents without rapport-building capabilities (Lucas, Boberg, et al., 2018a, 2018b). However, this only seems to occur if rapport-building occurs before the persuasive conversation (but not after; Lucas, Boberg, et al., 2018b) and only among participants acculturated in the U.S. (Lucas, Boberg, et al., 2018a). Additionally, we did not find that people were more comfortable with agents than human-operated agents in a personal training context (Lucas, Krämer, et al., 2018). While we expected participants using our virtual trainer to be concerned about being judged, like in the clinical and financial interview contexts, and thus respond more positively to the virtual intelligent agent than human-operated agents, they did not. In this work, there are also implications for rapport with virtual trainers: people felt more rapport when they believed the virtual trainer was an operated by a human than when they believed it was a virtual intelligent agent. Some studies in other contexts where people could be concerned about being judged, such as when practicing negotiation, have shown that people are more comfortable with virtual intelligent agents than human-operated agents (Gratch, DeVault, & Lucas, 2016), while others have not (Gratch, Devault, Lucas, & Marsella, 2015).

Characteristics of participants that impact rapport

The previous studies and findings emphasized the impact of differences in the agent's behavior but numerous social psychological studies emphasize that the trajectory of a social interaction is heavily shaped by the "baggage" people bring to a situation. In human-to-human interactions, people who are extroverted will more easily establish rapport than introverts and we might expect that these dispositional tendencies will carry over into interactions with rapport agents. Indeed, in a series of studies we have found that dispositional factors shape interactions with socially intelligent agents in similar ways to how they influence face-to-face interactions between people.

Indeed, several dispositional factors have been found to influence the effectiveness of rapport agents. With respect to the Big Five, extroversion and agreeableness influence interactions in ways that are consistent with their impact in human-to-human interactions. Extroverts tend to talk more, more fluently and feel better about their interaction and similar findings hold for participants that score high in agreeableness (Kang, Gratch, Wang, & Watts, 2008; von der Pütten et al., 2010). Social anxiety also plays a moderating role: interestingly, we found dispositionally anxious subjects felt more trust towards a rapport agent than in their interactions with human conversational partners (Kang, Gratch, Wang, & Watt, 2008).

Dispositional factors don't uniquely determine the outcome of a social interaction, but rather interact with aspects of the situation. For example, someone that is confident in social situations might perform well regardless of the behavior of their conversational partner. However, someone that is less secure might seek constant reassurance by carefully monitoring their partner's nonverbal feedback: if this feedback is positive, they may perform well and report positive feelings; if this feedback is negative, the opposite may occur. We see similar interaction effects with rapport agents. For example, extroverts seem insensitive to manipulations that impact the quality of a rapport agent's nonverbal feedback, whereas participants that score high in social anxiety are quite disrupted when they fail to receive positive and coordinated nonverbal feedback (Kang, Gratch, Wang, & Watt, 2008).

Overall, our studies suggest that both agent behavior and dispositional factors interact to determine the overall quality of an experience with a socially intelligent agent. Further, these effects are largely consistent with predictions from the social psychological literature on human-to-human interactions.

5. Discussion and Conclusion

Across a wide range of studies, we have consistently shown that simple nonverbal cues on the part of a computer program can provoke a wide range of beneficial subjective and behavioral outcomes. Overall, our studies and related findings provide substantial evidence that the nonverbal behavior of socially intelligent agents influence the behavior of the humans that interact with them in ways that are consistent with psychological findings on human-to-human interaction. Further, these effects increase as a virtual agent exhibits more human-like behavioral characteristics. More specifically, studies support

Tickle-Degnen and Rosenthal's claims that rapport is promoted by social behavioral cues that are positive, contingent and convey mutual attention, and that these effects are moderated by personality traits of the human user.

Despite the apparent success of rapport agents, one should be cautious before concluding that people will always be so easily manipulated by simple nonverbal behaviors. Behaviors such as nods or smiles might trigger automatic responses and simple social inferences but is otherwise limited. Our experimental settings (storytelling and interviews) are, by design, simple for agents to navigate. For example, the Rapport Agent conveys understanding without actual understanding, a behavior that most of us have engaged in from time to time (for example, when carrying on a conversation in a unfamiliar language or in a noisy room), but such a charade only goes so far before it ends in embarrassment. In a similar storytelling paradigm, Janet Bavelas illustrated that "generic" feedback (similar to what the Rapport Agent provides) is easy to produce without actually attending to the meaning of a conversation – she had participants listen while solving complex mathematical problems – but at certain points speakers need more meaningful "specific" feedback. Bavelas et al. (2000) found when speakers were telling personally emotional stories, they expected emotional feedback at key dramatic moments. When they failed to receive it, they felt embarrassment and had difficulty proceeding with their story. Even more generic feedback requires some level of natural language sophistication. According to theories of conversational grounding, speakers in a conversation expect frequent and incremental feedback from listeners that their communication is understood (Nakano, Reinstein, Stocky, & Cassell, 2003; Traum, 1994). When listeners provide grounding feedback, speech can proceed fluently and presumably with a greater sense of rapport. Such feedback often takes the form of nods, such as produced by our rapport agents, and presumably speakers are (incorrectly) interpreting these nods as grounding cues. This illusion can be maintained to an extent, but it will eventually backfire and lead participants to view such feedback with suspicion. SimSensei, with its incorporation of a dialog manager, does have the ability to provide some cognitive and emotional feedback, but in practice this had to be used sparingly due to recognition errors. Thus, the ability to provide rapid and meaningful feedback remains a challenge for rapport agents.

Part of our research can be seen as pushing the boundary of just how far one can go with simple contingent feedback. Our early studies explored "safe" and impersonal content such as cartoons. Over time we took on more challenging setting such as mental health screenings. At each stage, we continue to show robust subjective and behavioral effects of contingent positive feedback, however we must be careful before concluding that the agent is performing well. What we've shown is that the agent performs about as well as discussing personal matters with a stranger (something many people are uncomfortable with) and better than an agent that provides no or negative feedback. While important, this is a low bar and much can be done to improve the performance and effects of such agents. Future research must extend beyond such "mindless" feedback (i.e., feedback without deep understanding) to ensure that responses are aligned with the underlying grounding and inferential mechanisms of social agents. Indeed, research on emotional expressions emphasize that people assume these reflect an agent's or "appraisal" of the current social situation (de Melo, Carnevale, Read, & Gratch, 2014; Hareli & Hess, 2010). An agent that fails to make these expressions contingent on the meaning of conversation actions will certainly reduce rapport. Thus, future work must better align nonverbal behaviors with the underlying cognitive machinery of agents. Alas, this point has been long recognized by the intelligent virtual agent community (see Gratch et al., 2002), but stubbornly hard to achieve.

To conclude, we have illustrated that interpersonal processes can be effectively modeled in computational systems, both virtual and robotic, and these models can yield important theoretical insights and practical benefits. Herb Simon emphasized the importance of a partnership between computational and social science (Simon, 1969) and research into rapport agents vividly illustrates this partnership continues to bear fruit. By starting with a psychological theory of social processes (in this case the theory of rapport by Tickle-Degnen and Rosenthal), we are able to construct a social artifact that instantiates the theory. This results in practical benefits, such as greater disclosure in mental health interviews, but such artifacts also provide the means to experimentally tease apart these factors in ways that avoid the disadvantages and potential confounds introduced by more traditional psychological methods (such as the use of human confederates). Thus, social theory has allowed us to build a better computer and return the favor through experimental support for theory. This is a true partnership between the social and computational sciences of social emotions.

6. References

Abbe, A., & Brandon, S. E. (2013). The Role of Rapport in Investigative Interviewing: A Review. *Journal of Investigative Psychology and Offender Profiling, 10*(3), 237-249. doi:10.1002/jip.1386

Adam, H., Shirako, A., & Maddux, W. W. (2010). Cultural Variance in the Interpersonal Effects of Anger in Negotiations. *Psychological Science, 21*(6), 882-889. doi:10.1177/0956797610370755

Ambady, N., Bernieri, F. J., & Richeson, J. A. I. E., , vol. 32 (pp. 201–271) San Diego: Academic. (Eds.). (2000). *Toward a histology of social behavior: Judgment accuracy from thin slices of the behavioral stream* (Vol. 32).

Bailenson, J., Beall, A., Loomis, J., Blascovich, J., & Turk, M. (2004). Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments. *Presence: Teleoperators and Virtual Environments, 13*(4), 428-441.

Bailenson, J., & Yee, N. (2005). Digital Chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16*, 814-819.

Baker, R. P. (1992). New Technology in Survey Research: Computer-Assisted Personal Interviewing (CAPI). *Social Science Computer Review, 10*(2), 145-157. doi:10.1177/089443939201000202

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of Social Behavior: Direct Effects of Trait Construct and Stereotype Activation on Action. *Journal of Personality and Social Psychology, 71*(2), 230-244.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as Co-narrators. *Journal of Personality and Social Psychology, 79*(6), 941-952.

Beckenbach, A. (1995). Computer-Assisted Questioning: the New Survey Methods in the Perception of the Respondents. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 48*(1), 82-100. doi:10.1177/075910639504800111

Bente, G., Kraemer, N. C., Petersen, A., & de Ruiter, J. P. (2001). Computer Animated Movement and Person Perception: Methodological Advances in Nonverbal Behavior Research. *Journal of Nonverbal Behavior, 25*(3), 151-166.

Bergmann, K., & Kopp, S. (2012). *Gestural alignment in natural dialogue.* Paper presented at the Proceedings of the Annual Meeting of the Cognitive Science Society.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of nonverbal behavior, 12*(2), 120-138. doi:10.1007/BF00986930

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of Nonverbal Behavior*. Cambridge: Cambridge Univ. Press.

Biocca, F., & Harms, C. (2002). *Defining and measuring social presence: Contribution to the net-worked minds theory and measure*. Paper presented at the Presence.

Blascovich, J. (2014). Challenge, threat, and social influence in digital immersive virtual environments. *Social Emotions in Nature and Artefact. Oxford University Press, New York*, 44-54.

Borrie, S. A., Barrett, T. S., Willi, M. M., & Berisha, V. (2019). Syncing Up for a Good Conversation: A Clinically Meaningful Methodology for Capturing Conversational Entrainment in the Speech Domain. *Journal of Speech, Language, and Hearing Research, 62*(2), 283-296. doi:doi:10.1044/2018_JSLHR-S-18-0210

Brennan, S. E. (1996). Lexical entrainment in spontaneous dialog. *Proceedings of ISSD, 96*, 41-44.

Burgoon, J. K. (1983). Nonverbal violations of expectations. In J. M. Wiemann & R. P. Harrison (Eds.), *Nonverbal Interaction* (pp. 11-77). Beverly Hils, CA: Sage.

Burgoon, J. K., Guerrero, L. K., & Manusov, V. (2016). *Nonverbal communication*: Routledge.

Burgoon, J. K., Stern, L. A., & Dillman, L. (1995). *Interpersonal adaptation: Dyadic interaction patterns*. Cambridge: Cmabridge University Press.

Burns, M. (1984). Rapport and relationships: The basis of child care. *Journal of Child Care, 2*, 47-57.

Cassell, J., Gill, A., & Tepper, P. (2007). *Conversational Coordination and Rapport.* Paper presented at the Proceedings of Workshop on Embodied Language Processing at ACL 2007, Prague, CZ.

Charny, J. E. (1966). Psychosomatic manifestations of rapport in psychotherapy. *Psychosomatic medicine, 28*(4), 305-315.

Chartrand, T. L., & Bargh, J. A. (1999). The Chameleon Effect: The Perception-Behavior Link and Social Interaction. *Journal of Personality and Social Psychology, 76*(6), 893-910.

Coates, J. (1989). Gossip revisited: Language in all-female groups. *Women in their speech communities*, 94-122.

Cogger, J. W. (1982). Are you a skilled interviewer? *Personnel Journal, 61*, 840-843.

de Kok, I., & Heylen, D. (2011). The MultiLis corpus–Dealing with individual differences in nonverbal listening behavior *Toward Autonomous, Adaptive, and Context-Aware Multimodal Interfaces. Theoretical and Practical Issues* (pp. 362-375): Springer.

de Melo, C., Carnevale, P. J., & Gratch, J. (2014). *Using virtual condederates to research intergroup bias and conflict*. Paper presented at the 74th Annual Meeting of the Academy of Management, Philadelphia, PA.

de Melo, C., Carnevale, P. J., Read, S. J., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology, 106*(1), 73-88. doi:10.1037/a0034251

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., . . . Morency, L.-P. (2014). *SimSensei kiosk: a virtual human interviewer for healthcare decision support*. Paper presented at the Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems, Paris, France.

DeVault, D., Sagae, K., & Traum, D. (2011). Incremental interpretation and prediction of utterance meaning for interactive dialogue. *Dialogue & Discourse, 2*(1), 143-170.

Dijkstra, W. (1987). Interviewing Style and Respondent Behavior:An Experimental Study of the Survey-Interview. *Sociological Methods & Research, 16*(2), 309-334. doi:10.1177/0049124187016002006

Drolet, A. L., & Morris, M. W. (2000). Rapport in conflict resolution: accounting for how face-to-face contact fosters mutual cooperation in mixed-motive conflicts. *Experimental Social Psychology, 36*, 26-50.

Duncan, S., Franklin, A., Parrill, F., Welji, H., Kimbara, I., & Webb, R. (2007). *Cognitive processing effects of 'social resonance 'in interaction.* Paper presented at the In: Proceedings Gesture 2007-the conference of the Int. Society of Gesture Studies.

Forgas, J. P. (2007). The use of computer-mediated interaction in exploring affective influences on strategic interpersonal behaviours. *Computers in Human Behavior, 23*(2), 901-919.

Fredrickson, B. L. (2016). Love: Positivity resonance as a fresh, evidence-based perspective on an age-old topic. *Handbook of emotions*, 847-858.

Friedberg, H., Litman, D., & Paletz, S. B. F. (2012, 2-5 Dec. 2012). *Lexical entrainment and success in student engineering groups.* Paper presented at the 2012 IEEE Spoken Language Technology Workshop (SLT).

Fuchs, D. (1987). Examiner familiarity effects on test performance: implications for training and practice. *Topics in Early Childhood Special Education, 7*, 90-104.

Fujie, S., Ejiri, Y., Nakajima, K., Matsusaka, Y., & Kobayashi, T. (2004). *A Conversation Robot Using Head Gesture Recognition as Para-Linguistic Information.* Paper presented at the 13th IEEE International Workshop on Robot and Human Communication.

Goldberg, S. B. (2005). The secrets of successful mediators. *Negotiation Journal, 21*(3), 365-376.

Gratch, J., DeVault, D., & Lucas, G. (2016). *The benefits of virtual humans for teaching negotiation*. Paper presented at the 16th International Conference on Intelligent Virtual Agents, Los Angeles, CA.

Gratch, J., Devault, D., Lucas, G., & Marsella, S. (2015). *Negotiation as a Challenge Problem for Virtual Humans*. Paper presented at the 15th International Conference on Intelligent Virtual Agents, Delft, The Netherlands.

Gratch, J., Kang, S.-H., & Wang, N. (2013). Using social agents to explore theories of rapport and emotional resonance In J. Gratch & S. Marsella (Eds.), *Social Emotions in Nature and Artifact*: Oxford University Press.

Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., van der Werf, R., & Morency, L.-P. (2006). *Virtual Rapport.* Paper presented at the 6th International Conference on Intelligent Virtual Agents, Marina del Rey, CA.

Gratch, J., Rickel, J., André, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating Interactive Virtual Humans: Some Assembly Required. *IEEE Intelligent Systems, July/August,* 54-61.

Gratch, J., Wang, N., Gerten, J., & Fast, E. (2007). *Creating Rapport with Virtual Agents*. Paper presented at the 7th International Conference on Intelligent Virtual Agents, Paris, France.

Gremler, D. D., & Gwinner, K. P. (2000). Customer-Employee Rapport in Service Relationships. *Journal of Service Research, 3*(1), 82-104. doi:10.1177/109467050031006

Hall, J. A., Harrigan, J. A., & Rosenthal, R. (1995). Nonverbal behavior in clinician—patient interaction. *Applied and Preventive Psychology, 4*(1), 21-37. doi:10.1016/s0962-1849(05)80049-6

Hareli, S., & Hess, U. (2010). What emotional reactions can tell us about the nature of others: An appraisal perspective on person perception. *Cognition and Emotion, 24*(1), 128-140.

Hoegen, R., Schalk, J. v. d., Lucas, G., & Gratch, J. (2018). *The impact of agent facial mimicry on social behavior in a prisoner's dilemma*. Paper presented at the Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, NSW, Australia.

Horton, D., & Wohl, R. R. (1956). Mass Communication and Para-social Interaction: Observations on Intimacy at a Distance. *Psychiatry, 19*(215-29).

Huang, L. (2013). *Parasocial Consensus Sampling: Modeling Human Nonverbal Behaviors from Multiple Perspectives.* University of Southern California.

Huang, L., & Gratch, J. (2013). *Explaining the Variability of Human Nonverbal Behaviors in Face-to-Face Interaction.* Paper presented at the International Workshop on Intelligent Virtual Agents.

Huang, L., Morency, L.-P., & Gratch, J. (2010). Learning Backchannel Prediction Model from Parasocial Consensus Sampling: A Subjective Evaluation. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Intelligent Virtual Agents* (Vol. 6356, pp. 159-172): Springer Berlin Heidelberg.

Huang, L., Morency, L.-P., & Gratch, J. (2011). *Virtual rapport 2.0*. Paper presented at the Proceedings of the 10th international conference on Intelligent virtual agents, Reykjavik, Iceland.

Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity (Vol. 31, pp. 177-192).

Julien, D., Brault, M., Chartrand, É., & Bégin, J. (2000). Immediacy behaviours and synchrony in satisfied and dissatisfied couples. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 32*(2), 84.

Kallmann, M., & Marsella, S. (2005). *Hierarchical Motion Controllers for Real-Time Autonomous Virtual Humans.* Paper presented at the 5th International Working Conference on Intelligent Virtual Agents, Kos, Greece.

Kang, O., & Wheatley, T. (2017). Pupil dilation patterns spontaneously synchronize across individuals during shared attention. *Journal of Experimental Psychology: General, 146*(4), 569-576. doi:10.1037/xge0000271

Kang, S.-h., Gratch, J., Wang, N., & Watt, J. (2008). *Does Contingency of Agents' Nonverbal Feedback Affect Users' Social Anxiety?* Paper presented at the 7th International Conference on Autonomous Agents and Multiagent Systems, Estoril, Portugal.

Kang, S.-H., Gratch, J., Wang, N., & Watts, J. (2008). *Agreeable People Like Agreeable Virtual Humans.* Paper presented at the 8th International Conference on Intelligent Virtual Agents, Tokyo.

Karacora, B., Dehghani, M., Krämer, N. C., & Gratch, J. (2012). *The influence of virtual agents' gender and rapport on enhancing math performance.* Paper presented at the Proceedings of the Annual Meeting of the Cognitive Science Society.

Khooshabeh, P., & Lucas, G. (2018). Virtual Human Role Players for Studying Social Factors in Organizational Decision Making. *Frontiers in Psychology, 9*(194). doi:10.3389/fpsyg.2018.00194

Kopp, S. (2010). Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication, 52*(6), 587-597. doi:https://doi.org/10.1016/j.specom.2010.02.007

Kopp, S., Krenn, B., Marsella, S., Marshall, A., Pelachaud, C., Pirker, H., . . . Vilhjálmsson, H. (2006). *Towards a common framework for multimodal generation in ECAs: The behavior markup language.* Paper presented at the Intelligent Virtual Agents, Marina del Rey, CA.

Krämer, N. C., Karacora, B., Lucas, G., Dehghani, M., Rüther, G., & Gratch, J. (2016). Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education, 99*, 1-13. doi:https://doi.org/10.1016/j.compedu.2016.04.002

Krämer, N. C., Lucas, G., Schmitt, L., & Gratch, J. (2018). Social snacking with a virtual agent – On the interrelation of need to belong and effects of social responsiveness when interacting with artificial entities. *International Journal of Human-Computer Studies, 109*(Supplement C), 112-121. doi:https://doi.org/10.1016/j.ijhcs.2017.09.001

Lafferty, J., McCallum, A., & Pereira, F. C. (2001). Conditional random fields: Probabilistic models for segmenting and labeling sequence data.

Lakin, J. L., Jefferis, V. A., Cheng, C. M., & Chartrand, T. L. (2003). Chameleon Effect as Social Glue: Evidence for the Evolutionary Significance of Nonconsious Mimicry. *Journal of nonverbal behavior, 27*(3), 145-162.

Lanzetta, J. T., & Englis, B. G. (1989). Expectations of cooperation and competition and their effects on observers' vicarious emotional responses. *Journal of Personality and Social Psychology, 45*, 543-554.

Levitan, R., Gravano, A., & Hirschberg, J. (2011). *Entrainment in speech preceding backchannels.* Paper presented at the Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2.

Levitan, R., & Hirschberg, J. (2011). *Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions.* Paper presented at the Twelfth Annual Conference of the International Speech Communication Association.

Lubold, N., Walker, E., Pon-Barry, H., & Ogan, A. (2018, 2018//). *Automated Pitch Convergence Improves Learning in a Social, Teachable Robot for Middle School Mathematics.* Paper presented at the Artificial Intelligence in Education, Cham.

Lucas, G., Boberg, J., Traum, D., Artstein, R., Gratch, J., Gainer, A., . . . Nakano, M. (2018a). *Culture, Errors, and Rapport-building Dialogue in Social Agents*. Paper presented at the Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, NSW, Australia.

Lucas, G., Boberg, J., Traum, D., Artstein, R., Gratch, J., Gainer, A., . . . Nakano, M. (2018b). *Getting to Know Each Other: The Role of Social Dialogue in Recovery from Errors in Social Robots*. Paper presented at the Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, Chicago, IL, USA.

Lucas, G., Gratch, J., King, A., & Morency, L.-P. (2014). It’s only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94-100.

Lucas, G., Krämer, N. C., Peters, C., Taesch, L.-S., Mell, J., & Gratch, J. (2018). *Effects of Perceived Agency and Message Tone in Responding to a Virtual Personal Trainer*. Paper presented at the Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, NSW, Australia.

Lucas, G., Lehr, J., Krämer , N. C., & Gratch, J. (2019). *The Effectiveness of Social Influence Tactics when Used by a Virtual Agent*. Paper presented at the Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents, Paris, France.

Lucas, G., Rizzo, A., Gratch, J., Scherer, S., Stratou, G., Boberg, J., & Morency, L.-P. (2017). Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. *Frontiers in Robotics and AI, 4*(51). doi:10.3389/frobt.2017.00051

Marsella, S., Xu, Y., Lhommet, M., Feng, A., Scherer, S., & Shapiro, A. (2013). *Virtual character performance from speech.* Paper presented at the Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation.

McFarland, D. H. (2001). Respiratory Markers of Conversational Interaction. *Journal of Speech, Language, and Hearing Research, 44*, 128-143.

Mead, G. H. (1934). *Mind, self and society*

Chicago: University of Chicago Press.

Mell, J., Lucas, G., & Gratch, J. (2017). *Prestige Questions, Online Agents, and Gender-Driven Differences in Disclosure*, Cham.

Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology, 44*(6), 1234-1244. doi:10.1037/0022-3514.44.6.1234

Morbini, F., DeVault, D., Sagae, K., Gerten, J., Nazarian, A., & Traum, D. (2014). FLoReS: a forward looking, reward seeking, dialogue manager *Natural Interaction with Robots, Knowbots and Smartphones* (pp. 313-325): Springer.

Morency, L.-P., de Kok, I., & Gratch, J. (2008). *Predicting Listener Backchannels: A Probabilistic Multimodal Approach.* Paper presented at the 8th International Conference on Intelligent Virtual Agents, Tokyo.

Morency, L.-P., de Kok, I., & Gratch, J. (2010). A probabilistic multimodal approach for predicting listener backchannels. *Autonomous Agents and Multi-Agent Systems, 20*(1), 70-84. doi:10.1007/s10458-009-9092-y

Morency, L.-P., Sidner, C., Lee, C., & Darrell, T. (2005). *Contextual Recognition of Head Gestures.* Paper presented at the 7th International Conference on Multimodal Interactions, Torento, Italy.

Nakano, Y., Reinstein, G., Stocky, T., & Cassell, J. (2003). *Towards a Model of Face-to-Face Grounding.* Paper presented at the Annual Meeting of the Association for Computational Linguistics, Sapporo, Japan.

Parkinson, B. (2013). Processes of Emotional Meaning and Response Coordination. *Social Emotions in Nature and Artifact*, 29.

Pickard, M. D., Roster, C. A., & Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior, 65*, 23-30. doi:https://doi.org/10.1016/j.chb.2016.08.004

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.

Roloff, M. E., & Campion, D. E. (1987). Communication and reciprocity within intimate relationships. In M. E. Roloff & G. R. Miller (Eds.), *Interpersonal processes: New directions in communication research* (pp. 11-38). Beverly Hills, CA: Sage.

Sebestik, J., Zelon, H., DeWitt, D., O'Reilly, J., & McGowan, K. (1988). *Initial experiences with CAPI.* Paper presented at the Proceedings of the Bureau of the Census Fourth Annual Research Conference.

Simon, H. (1969). *The sciences of the Artificial*. Cam bridge, MA: MIT Press.

Slack, W. V., & Van Cura, L. J. (1968). Patient reaction to computer-based medical interviewing. *Computers and Biomedical Research, 1*(5), 527-531.

Thiebaux, M., Marshall, A., Marsella, S., & Kallmann, M. (2008). *SmartBody: Behavior Realization for Embodied Conversational Agents*. Paper presented at the International Conference on Autonomous Agents and Multi-Agent Systems, Portugal.

Thornberry, O., Rowe, B., & Biggar, R. (1991). Use of Capi With the U.S. National Health Interview Survey. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique, 30*(1), 27-43. doi:10.1177/075910639103000103

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and its Nonverbal Correlates. *Psychological Inquiry, 1*(4), 285-293.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology, 84*(3), 558-568.

Traum, D. (1994). *A computational theory of grouding in natural language conversation.* (Ph.D.), University of Rochester, Rochester, NY.

Truong, K. P., Poppe, R., & Heylen, D. (2010). *A rule-based backchannel prediction model using pitch and pause information.* Paper presented at the Eleventh Annual Conference of the International Speech Communication Association.

Tsui, P., & Schultz, G. L. (1985). Failure of Rapport: Why psychotheraputic engagement fails in the treatment of Asian clients. *American Journal of Orthopsychiatry, 55*, 561-569.

von der Pütten, A., Krämer, N. C., & Gratch, J. (2009). *Who´s there? Can a Virtual Agent Really Elicit Social Presence*. Paper presented at the 12th Annual International Workshop on Presence, Los Angeles.

von der Pütten, A., Krämer, N. C., & Gratch, J. (2010). How Our Personality Shapes Our Interactions with Virtual Characters - Implications for Research and Development. In J. Allbeck, N. Badler, T.

Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Intelligent Virtual Agents* (Vol. 6356, pp. 208-221): Springer Berlin / Heidelberg.

von der Pütten, A., Krämer, N. C., Gratch, J., & Kang, S.-H. (2009). *It doesn't matter what you are! Comparing interacting with an autonomous virtual person with interacting with a virtually represented human*. Paper presented at the Proceedings of the 6th Conference of the Media Psychology Division of the German Psychological Society, Lengerich.

Wang, N., & Gratch, J. (2009a). *Can a Virtual Human Build Rapport and Promote Learning?* Paper presented at the 14th International Conference on Artificial Intelligence in Education, Brighton.

Wang, N., & Gratch, J. (2009b). *Rapport and Facial Expression*. Paper presented at the International Conference on Affective Computing and Intelligent Interaction, Amsterdam.

Wang, N., & Gratch, J. (2010). *Don't just stare at me*. Paper presented at the 28th Annual CHI Conference on Human Factors in Computing Systems, Atlanta.

Ward, N., & Tsukahara, W. (2000). Prosodic features which cue back-channel responses in English and Japanese. *Journal of Pragmatics, 23*, 1177-1207.

Warner, C. H., Appenzeller, G. N., Grieger, T., Belenkiy, S., Breitbach, J., Parker, J., . . . Hoge, C. (2011). Importance of Anonymity to Encourage Honest Reporting in Mental Health Screening After Combat Deployment. *Archives of General Psychiatry, 68*(10), 1065-1071. doi:10.1001/archgenpsychiatry.2011.112

Weisband, S., & Kiesler, S. (1996). *Self disclosure on computer forms: meta-analysis and implcations*. Paper presented at the CHI.

Welji, H., & Duncan, S. (2004). *Characteristics of face-to-face interactions, with and without rapport: Friends vs. strangers.* Paper presented at the Symposium on Cognitive Processing Effects of 'Social Resonance' in Interaction, 26th Annual Meeting of the Cognitive Science Society.

Whitehill, J., Bartlett, M. S., & Movellan, J. R. (2008). *Automatic Facial Expression Recognition for Intelligent Tutoring Systems*. Paper presented at the CVPR 2008 Workshop on Human Communicative Behavior Analysis, Anchorage, Alaska.

Yngve, V. H. (1970). *On getting a word in edgewise.* Paper presented at the Sixth regional Meeting of the Chicago Linguistic Society.