First Impressions in Human–Agent Virtual Encounters

ANGELO CAFARO and HANNES HÖGNI VILHJÁLMSSON, Reykjavik University
TIMOTHY BICKMORE, Northeastern University

In greeting encounters, first impressions of personality and attitude are quickly formed and might determine important relational decisions, such as the likelihood and frequency of subsequent encounters. An anthropomorphic user interface is not immune to these judgments, specifically when exhibiting social interaction skills in public spaces. A favorable impression may help engaging users in interaction and attaining acceptance for long-term interactions. We present three studies implementing a model of first impressions for initiating user interactions with an anthropomorphic museum guide agent with socio-relational skills. We focus on nonverbal behavior exhibiting personality and interpersonal attitude. In two laboratory studies, we demonstrate that impressions of an agent's personality are quickly formed based on proximity, whereas interpersonal attitude is conveyed through smile and gaze. We also found that interpersonal attitude has greater impact than personality on the user's decision to spend time with the agent. These findings are then applied to a museum guide agent exhibited at the Boston Museum of Science. In this field study, we show that employing our model increases the number of visitors engaging in interaction.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/methodology*; I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Intelligent agents*; I.6.8 [**Simulation and Modeling**]: Types of Simulation—*Animation*; J.4 [**Computer Applications**]: Social and Behavioral Sciences—*Psychology sociology*

General Terms: Design, Experimentation, Human factors

Additional Key Words and Phrases: First impressions, nonverbal behavior, interpersonal attitudes, personality traits, relational agents

ACM Reference Format:
Angelo Cafaro, Hannes Högni Vilhjálmsson, and Timothy Bickmore. 2016. First impressions in human–agent virtual encounters. ACM Trans. Comput.-Hum. Interact. 23, 4, Article 24 (August 2016), 40 pages.
DOI: http://dx.doi.org/10.1145/2940325

24

1. INTRODUCTION

Every interaction we have with a new person entails a first impression [Ambady and Skowronski 2008]. The judgments we form of others based on a brief first encounter often have enormous staying power, with our initial perceptions continuing to be influential months later [Miller et al. 2007; Sunnafrank and Ramirez 2004] and determining

This work was conducted at the Center for Analysis and Design of Intelligent Agents with support from the School of Computer Science at Reykjavik University and the Icelandic Research Fund under Grant No. 130705051. Further support provided by the U.S. National Science Foundation under Grant No. 1012086.
Authors' addresses: A. Cafaro, CNRS-LTCI, Télécom ParisTech, Multimedia Group, Signal and Image Processing Department (TSI), 46 Rue Barrault, 75013 Paris, France; email: angelo.cafaro@telecom-paristech.fr; H. H. Vilhjálmsson, Socially Expressive Computing Group, School of Computer Science, Center for Analysis and Design of Intelligent Agents (CADIA), Reykjavik University, Menntavegur 1, 101 Reykjavik Iceland; email: hannes@ru.is; T. W. Bickmore, College of Computer and Information Science, Northeastern University, 202 West Village H, 360 Huntington Avenue, 02115 Boston, MA; email: bickmore@ccs.neu.edu.

DOI: http://dx.doi.org/10.1145/2940325

important relational decisions such as the likelihood and frequency of subsequent encounters [Riggio and Friedman 1986].

A software user interface is not immune to judgments [Saadé and Otrakji 2007], particularly, when embodied in a humanoid animated agent capable of engaging users in real-time multimodal interaction [Maat et al. 2009, 2010; Cafaro et al. 2012]. Humans engage with virtual agents and respond to their gestures and statements [Nunamaker et al. 2011]. These agents could be used as interfaces between users and computers [Deng et al. 2006], but a wrong (or unwanted) impression on users might jeopardize the achievement of important design goals, such as maintaining *engagement* starting from the initial moments of interaction and fostering technology *acceptance* needed for long-term interactions [Davis 1993].

In some applications of these agents deployed in public spaces, such as information kiosks [Cassell et al. 2002; Nunamaker et al. 2011], office companions [Deshmukh et al. 2010], or museum exhibits [Kopp et al. 2005; Bickmore et al. 2008; Al Moubayed et al. 2012; Traum et al. 2012], a multitude of users can potentially engage in an interaction with the agent. Therefore, welcoming and engaging the user by making a favorable impression is a fundamental design goal. In *museum exhibit applications*, for example, if visitors form negative judgments of the agent they are unlikely to engage in an interaction, or they might end it midway [Ghosh and Kuzuoka 2014]. On the other hand, technology acceptance is a critical requirement for interface agents supporting long-term interactions. Agents, designed for building long-term relationships, are known as *relational agents* (RAs) [Bickmore and Picard 2005], and previous research has demonstrated that being able to build a socio-emotional relationship with their users improves task outcomes in a variety of application domains including coaching [Bickmore et al. 2005, 2012], counseling [Kang et al. 2012], psychotherapy [Rizzo et al. 2013], and health care [Lucas et al. 2014; Zhou et al. 2014; Bickmore et al. 2009, 2010]. No matter whether the goal is short-term user engagement or acceptance for the long term, it cannot be achieved if users outright reject the agent during their first encounter.

In human–human interaction, first impressions can be formed by observing individual characteristics, such as height, clothing or, generally, visual appearance [Naumann et al. 2009; Argyle 1988; Miller et al. 2007], from stereotypes [Bar et al. 2006; Ambady and Skowronski 2008], but also from behavioral characteristics, such as facial expressions and body language (i.e., nonverbal behavior) [Riggio and Friedman 1986; Argyle 1988; Burgoon et al. 1984]. In the context of first encounters, individuals can carefully plan how to present themselves visually, but it may be difficult to have full control over all nonverbal cues during the interaction [DePaulo 1992]. After just a few seconds of observing someone's nonverbal behavior we infer, with remarkable accuracy, a variety of information including, for instance, a person's skills [Ambady and Rosenthal 1993], sexual orientation [Rule and Ambady 2008], political view [Rule and Ambady 2010b], personality and interpersonal attitude [Argyle 1988; Levesque and Kenny 1993; Richmond et al. 2008; Campbell and Rushton 1978].

This would suggest that during a first human–agent interaction it might be important to make the appropriate choice of nonverbal behavior exhibited by an agent. The context where the interaction takes place might further determine the impressions that are more easily recognized, possibly due to expectations that a person brings to the interaction [Ambady and Skowronski 2008]. In zero-acquaintance situations, such as for a guide welcoming visitors at a museum, the behaviors can help frame the interaction using a friendly attitude and portray an outgoing, gregarious, and sociable personality, resulting in a more favorable impression [Richmond et al. 2008].

Framing the interaction with a specific *interpersonal attitude* toward the user or portraying a certain *personality* can serve an important purpose: It gives the user a glimpse of how the subsequent interactions with the agent could be. Incorporating these dimensions has been recommended for the design of intelligent interfaces for museum visits [Stock and Zancanaro 2007]. Appropriate social behavior for providing museum guides with personality is needed [Richards 2012]. However, there are several challenges to consider when attempting to concurrently manage those impressions through nonverbal behavior. First, an attitude is expressed toward a particular person and it can dynamically evolve (along a positive/negative valence) during the interaction based on how the interactant responds [Argyle 1988]. Personality remains more stable over time and does not depend on the particular interactant [Miller et al. 2007]. Furthermore, behaviors that typically occur in greeting interactions, such as gaze and proxemics behavior [Kendon 1990], may account for both dimensions and not necessarily yield the desired outcome. Finally, controlling for both aspects might become hard when behaviors are exhibited in small but still perceivable doses. For instance, consider a real-life first encounter where you are expecting to meet and welcome someone and aim at showing friendliness, but at the same time you attempt to not appear too extroverted. You might ask yourself: "Could looking at the person too much be judged as unfriendly?", also "Would it make sense to step toward the person or should I wait for them?".

In this paper, we present a first impression model for initiating user–agent interactions. This model extends conversational agent interaction models by taking into account the first phases of interaction (i.e., the greeting) and considers the impressions that users might form during those phases. We describe three empirical studies implementing this model in which we manipulate a museum agent's nonverbal immediacy cues (i.e., smile, gaze, and proxemic behavior) during the initial phases of a greeting encounter with the user, first in a virtual environment, and then in a real public space (i.e., Boston Museum of Science (MoS)). We focus on behavior for managing personality and interpersonal attitude impressions, as they are easily recognizable in such short interactions, but yet important factors determining engagement with the agent and user's acceptance of it. The museum context offers the opportunity to study first encounters in a setting in which (1) making favorable impressions is important, (2) first encounters (with visitors) are very likely to happen, (3) controlled experiments are still possible and easier to conduct (compared to outdoor locations), and (4) a great number of participants can be easily recruited in a natural and unbiased setting. Moreover, there is a growing interest in making conversational agents capable of interacting with uninformed humans in everyday situations [Kopp et al. 2005; Bickmore et al. 2008; Al Moubayed et al. 2012], making them step out of the laboratory and into real-world settings [Deshmukh et al. 2010].

In two initial laboratory preparation studies, we first wanted to understand what impressions of the agent's personality and interpersonal attitude the users form based on observed subtle changes of nonverbal behavior. We then wanted to study the effect of first impressions on users' willingness to interact with an agent again. Finally, we studied the effects of first impressions on user willingness to engage in interaction with a public space deployment of an RA.

The approach proposed in this paper focuses on the importance of first impressions in terms of user's acceptance and propensity to have subsequent interactions with an RA. However, it can be extended to conversational agents deployed in similar social contexts, both in the virtual and the real world, where the interaction can be initiated from distance. We apply psychosocial principles of human–human interaction to human–agent interaction, but our focus is not on the user's mental representation of the agent. Von der Pütten and colleagues [2010] suggest that it does not make a difference for either social reactions or social evaluations whether people believe they are interacting with another person or an artificial entity. Therefore, regardless of whether in the user's mind the agent is considered an artificial (i.e., artifact) or a real social entity, our focus is on the produced agent output (i.e., human-based nonverbal behavior), the user's interpretation of that behavior in terms of first impression judgments, and the impact on user's acceptance and willingness to interact with the agent.

2. NONVERBAL BEHAVIOR IN ZERO-ACQUAINTANCE HUMAN GREETINGS

The scenario we are focusing on is a greeting encounter between strangers, also known as a *zero-acquaintance* situation in the social psychology literature [Ambady and Skowronski 2008, p. 129]. Greetings have an important function in the management of relations, for example to begin, confirm, and continue a friendship [Kendon 1990, p. 154], and every encounter is unique with the many possible variations of the interactants' exhibited nonvebral behavior in response to each other socio-emotional cues [Katz 2001]. In particular, nonverbal cues during a greeting are powerful signals [Argyle 1988]. Smiling is a common feature of greeting rituals indicating the desire to interact [Kendon 1990, p. 189]. Spatial moves (proxemics) are also very powerful means for starting and ending an encounter [Argyle 1988], though they can be accompanied by other nonverbal signals, such as looking [Kendon 1990, p. 189]. Spatial behavior may also signal a particular definition of the situation. Just as orientation communicates co-operative and competitive relationships, so does greater distance indicate desire for greater formality [Argyle 1988].

Because of their presumed centrality to relational communication, the specific subset of cues identified by *smiling*, *gazing,* and *proxemics* is also categorized as "nonverbal immediacy behavior." *Immediacy* can be defined as the degree of perceived physical or psychological closeness between people [Richmond et al. 2008], and those behaviors can be valuable means for determining the personality or attitude of a communicator [Argyle 1988]. In a first greeting encounter, judgments of an individual's personality and attitude are of critical importance when considering starting a relationship [Miller et al. 2007] and for predicting future behavior [Borkenau et al. 2004]. Forming those impressions occurs spontaneously, without intention or even awareness [Ambady and Skowronski 2008, p. 108].

Personality refers to enduring characteristics of an individual that are revealed, for example, with patterns of exhibited nonverbal behavior [Richmond et al. 2008]. Personality researchers have identified a major approach to the study of human personality by measuring traits, which can be defined as broad themes in behavior, thoughts, and emotions that distinguish one person from another [Miller et al. 2007]. This approach assumes that behavior is determined by relatively stable traits, which are fundamental units of one's personality [Richmond et al. 2008]; these traits used to describe personality range from a few to, potentially, an unlimited number. According to Eysenck, for example, personality is reducible to three major traits: extraversion, neuroticism, and psychoticism [Eysenck 1953]. Cattell [1965] argued that it is was necessary to look at a much larger number of traits in order to get a complete picture of someone's personality, thus proposing a model of 16 traits.

Concerted efforts at factors analysis, however, have demonstrated that five factors are sufficient for providing the best compromise between explanatory power and parsimony. These are the Big 5 personality traits [McCrae and Costa 1997]: extraversion, neuroticism, agreeableness, conscientiousness, and openness to experience. Some of these traits are more visible than others [Miller et al. 2007, p. 140–141]. *Extraversion* (the extent to which people are outgoing, gregarious, talkative, and sociable), in particular, seems to be one of the easiest traits to pick up [Kammrath et al. 2007] through rapid interpersonal judgments of nonverbal behavior [Campbell and Rushton 1978], including interpersonal distance, smile, gaze, and posture [Argyle 1988; Burgoon et al. 1984; Richmond et al. 2008]. Moreover, positive impressions of extraversion compared to the other traits are more stable and they resist in the long term, which means they are not likely to be shaken by occasional observations of quiet or introverted behavior [Kammrath et al. 2007]. Someone's extraversion level could represent crucial information to answer adaptive questions such as whom to mate with or rely on in social alliances [Mehu et al. 2008]. We adopted the Big 5 theory in our first impression model (described in Section 3). The focused trait (i.e., extraversion) is common to the trait models discussed; however, the choice of five traits (compared to 16 as proposed by Cattell) simplified the design of our empirical studies (described in Section 5), and in particular, allowed us to assess the agent's and user's personality traits with a quicker but more effective methodology [Saucier 1994]. Moreover, the Big 5 model has been demonstrated to be effective in other conversational agents' research studies [Isbister and Nass 2000; Sevin et al. 2010; Neff et al. 2011; Liu et al. 2013; Krenn et al. 2014].

Compared to personality, interpersonal attitudes are expressed as a function of the person we are interacting with and are subject to a greater degree of variation. We can express different attitudes with different interactants, and during an interaction with an individual, we can shift our attitude, for example, from being friendly to unfriendly (or vice-versa) [Miller et al. 2007]. In sum, interpersonal attitudes are essentially an individual's conscious or unconscious evaluation of how s/he feels about and relates to another person [Argyle 1988, p. 85].

Several researchers attempted to identify the dimensions that can best represent the different interpersonal attitudes expressed during social interaction. Schutz [1958] proposed the dimensions of inclusion, control, and affect. Burgoon and Hale [1984] identified 12 dimensions defining different communication styles (e.g., dominance, intimacy, affect, engagement, inclusion, and confidence). Argyle proposed a two-dimensional representation. A first dimension is the *affiliation*, characterized as the degree of liking or friendliness and ranging from unfriendly to friendly. A second dimension is the *status*, related to power and assertiveness during the interaction and ranging from submissive to dominant [Argyle 1988, p. 86]. Affiliation, in particular, is broadly characterized as liking or wanting a close relationship. Thus, controlling affiliation in a first encounter, i.e., being friendly, can help to establish a relationship with another person [Argyle 1988, p. 86].

We are using Argyle's model to represent interpersonal attitudes. More specifically, as a first step, we are focusing on the affiliation dimension (unfriendly/friendly). This model has been successfully exploited in previous agents work [Lee and Marsella 2011; Ravenet et al. 2013; Chollet et al. 2014].

Those impressions of personality and attitude based on observable nonverbal behavior can be surprisingly accurate [Ambady and Skowronski 2008, p. 111]. They can be powerful determinants of how we behave toward another person [Rule and Ambady 2010a] and, most importantly, they shape expectations that we bring to future encounters [Goffman 1959]. According to the *predicted outcome theory* [Sunnafrank and Ramirez 2004], in social environments where potential relational partners will be proximate to one another in the future, individuals assess the likely outcomes of a future relationship to determine whether to develop the relationship and, if so, what type of relationship to attempt and how to proceed [Sunnafrank 1986].

In the specific social context of greeting encounters, there are precise social norms that regulate the interaction between participants [Kendon 1990], and nonverbal behavior is exhibited in relation to the physical space that separates them [Hall 1966]. Considering an encounter where a person is getting closer to another, the decreasing interpersonal distance becomes a trigger for certain nonverbal behavior that in turn shapes the first impressions [Argyle 1988].

Two important theories aim to clarify how humans interpret the space around them and behavior during greeting encounters. Hall [1966] described *proxemics*, the space around each individual, as consisting of four concentric areas that afford different kinds of communication with other individuals. From the closest to the furthest away, Hall labeled these areas *intimate*, *personal*, *social,* and *public*. Based on his theory, most social contact between acquaintances would occur within the social area, whereas the personal and intimate distances are reserved for very close contact with family and good friends.

Nonverbal behavior might not be exhibited in a discrete fashion associated within each proxemics zone, but it is likely to be part of an overall continuous process that spans across multiple zones. In fact how far one "goes out of one's way" to meet another appears to have a precise communicative significance [Kendon 1990, p. 179] as thoroughly described in Kendon's model of human greetings. According to Kendon's *greeting model* [1990, p. 153], a "greeting" is the unit of social interaction often observed when people come into one another's presence. Individuals that wish to greet each other adhere to a structured interaction comprising the following phases: *sighting*, *distant salutation*, *approach*, *close salutation,* and *initiation* of *conversation*. These phases come with an implicit relation to proxemics. In fact, specific behaviors are displayed by greeters at inexact but predictable distances. For example, a short exchange of glances as invitation cue to begin the greeting process happens within the sighting phase [Kendon 1990, p. 167] and a smile is often exhibited during the distant salutation as greeters continue to interact, move closer to one another afterward and commit to engage in interaction [Kendon 1990, p. 172].

In conclusion, greetings serve as precursors leading to interaction [Kendon 1990, p. 154], and how many stages of the greeting ritual are enacted will depend upon the kind of relationship that exists between the greeters. For a given kind of relationship, there are situations where all the steps of the ritual are appropriate, but others in which only some of them will be necessary. In situations where close interaction is to follow, for example, it appears that upon an initial encounter (i.e., the first time two individuals encounter one another within the context of a given occasion) both a distance and a close exchange salutation will occur, whereas upon subsequent encounters only a distant salutation is necessary [Kendon 1990, p. 203].

3. APPROACH: FIRST IMPRESSIONS MODEL

Figure 1 depicts our first impression model for initiating user–agent interactions. Our approach combines several theories, introduced earlier, for the creation of a model of user–agent interaction that considers the formative moments of the interaction (i.e., greeting phase) and the first impressions that users might form of an agent's personality and attitude in those phases. Based on Kendon's greeting model and Hall's proxemics theory, an RA should be able to encourage and engage the user in an interaction by exhibiting the appropriate nonverbal immediacy behavior in the very first steps of the greeting process (i.e., interaction).

The user–agent interpersonal distance, either in a virtual world or in the physical world, represents a valuable triggering mechanism for agent's nonverbal reactions as suggested by Kendon and Hall. On top of this, the *subtle* manipulations of agent's nonverbal immediacy cues of smile, gaze, and proxemics, in addition to serving basic greeting communicative functions, might also account for a user impressions of the agent's *extraversion* (personality trait according to the Big 5 model) and *affiliation* (the interpersonal attitude according to Argyle's model). These impressions may depend on the user's own personality. Then, according to the predicted outcome value theory, they may be predictive of future relational decisions, in terms of how likely it is that they would interact with the agent in the future and how often. User's personality traits can be predictors of the quality of an individual's relationships [Asendorpf and Wilpers 1998] and, interestingly, of tendency toward technology acceptance [Devaraj et al. 2008;

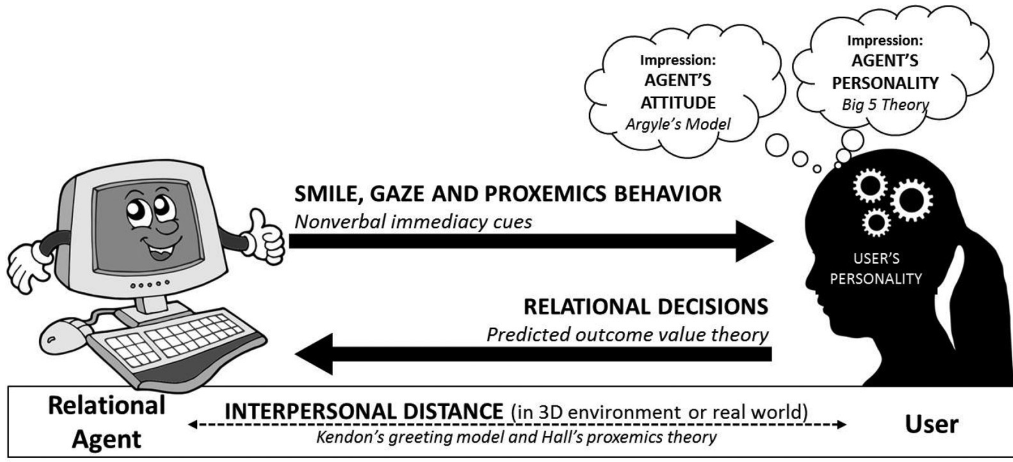

Fig. 1. A conceptual view of the combination of theories involved in our model of first impressions grounded in Kendon's, Hall's, and Argyle's models, and the Big 5 and predicted value outcome theories. As the interpersonal distance between user and agent decreases (either in a 3D environment or the real world), the nonverbal immediacy cues of smile, gaze, and proxemics exhibited by the agent are interpreted by the user and, depending also on his/her own personality, contribute to shaping impressions of the agent's personality and attitude. As outcome, these impressions impact users' relational decisions about further interactions with the agent, thus impacting the long-term acceptance of the agent.

Shengli et al. 2013; Barnett et al. 2015]. In our context of human–agent interaction, the two Big 5 factors that are most closely related to social interaction (i.e., extraversion and agreeableness) have been found to be predictive of positive agent's judgments in interpersonal encounters [Pütten et al. 2010]. Therefore, we included the assessment of user's personality in our model as it might impact the behavior interpretation when forming impressions of the agent, and the acceptance in terms of willingness to work with the agent.

This is a challenging approach that involves interesting research questions. First, the variety of nonverbal behavior exhibited by an agent might result in different user *interpretations* that are not necessarily obvious or might not mirror the findings in human social psychology. We have deliberately chosen nonverbal cues that (a) might account both for personality traits (i.e., extraversion) and social attitudes (i.e., affiliation), and (b) are subtly displayed. How are these cues interpreted when exhibited in orchestration by an agent? Does each cue account more for extraversion or friendliness impressions? Are small changes still perceived and having effects on users' first impressions? For example, would it be more convenient to gaze at the user and perform a little step toward him/her when s/he gets close? Would this be interpreted as a sign of extraversion and friendliness? Or would this be seen still as extraverted behavior, but unfriendly given context?

The second important aspect is the *impact* of the agent's exhibited behavior in terms of engagement and acceptance. First impressions of the agent's extraversion and affiliation at the same time may lead to unpredictable outcomes. Considering the different ways personality and attitudes endure over time, we aim to investigate how the two combine together and whether one of the two has greater impact on users' engagement and acceptance.

Finally, we explore these effects in a real-world setting. Our context is a greeting encounter in a museum where the agent has the role to greet and welcome a visitor.

When deploying virtual agents in such a *public space*, would laboratory findings still be applicable? Is the first impressions model effective?

We address these challenges with three user evaluation studies.

(1) *Nonverbal Behavior Interpretation Study.* We evaluate users' impressions of a greeting agent's extraversion and affiliation as a result of interpreting subtle nonverbal immediacy cues of smile, gaze, and proxemics in a first virtual encounter [Cafaro et al. 2012].
(2) *Nonverbal Behavior Impact Study.* We investigate whether the first impressions of extraversion and affiliation that users form of an RA have an impact on subsequent relational behavior of the user [Cafaro et al. 2013].
(3) *Managing First Impressions in a Public Space Study.* We move out of the virtual world to test the effectiveness of managing first impressions for an RA installed in a real public setting (i.e., Boston (MoS)), with thousands of people approaching it, therefore working with physical interpersonal distance between the user and the agent.

3.1. Study Design Principles

The first two studies are laboratory preparation studies aimed at accurately gathering the necessary information to answer our questions on interpretation and impact in a controlled setting. We considered presentation technology and agent's appearance (including embodiment, gender, and level or realism) as important design principles for creating the stimuli for our studies. These different choices were influenced both by theoretical and practical aspects.

Considering our ultimate deployment at the museum, we employed a virtual anthropomorphic robotic character for presenting the laboratory studies stimuli. The robotic appearance, as opposed to a more realistic appearance, allowed us to concentrate on the essential stimuli (i.e., nonverbal behavior), thus safely avoiding unwanted biases on users' impressions due to the clothing style. In addition, cartoon-style characters have been found to create lower behavioral realism expectations [McDonnell et al. 2012; Ring et al. 2014b] compared to more realistic characters.

4. RELATED WORK

Relational Agents. There are a significant number of studies on RAs, in different application domains. We survey here a few key findings. Cassell et al. [1999] created a virtual real-estate agent called REA that was capable of greeting her clients, finding out what sort of housing they were looking for, and finally, demonstrating a few matching properties. REA embraced a natural face-to-face conversation paradigm, with speech and gesture replacing the standard keyboard and mouse paradigm. Because it is important for real estate agents to build rapport (e.g., before asking a client sensitive questions about financial matters), Bickmore and Cassell [2001] added a small talk planning module and some immediacy behavior (mainly facial expressions). Later, based on the success of this idea, Bickmore and Picard [2005] introduced the concept of RAs.

In the health care domain, empathetic RAs provided discharge information to hospital patients with low health literacy [Bickmore et al. 2009] and with depressive symptoms [Bickmore et al. 2010]. The first RA used to promote anti-psychotic medication adherence among patients with schizophrenia is described in Bickmore et al. [2010]. In the counseling domain, it has been showed that an RA's nonverbal immediacy helps induce self-disclosure in human clients [Kang et al. 2012; Lucas et al. 2014] and supports longitudinal changes in user's health behavior [Ren et al. 2014].

Coaching is another successful application of RAs. Wang et al. [2009] investigated the effectiveness of RAs nonverbal immediacy behavior in creating rapport with a human learner and promoting learning. Bickmore et al. [2005] created an exercise advisor for older adults that was effective at promoting physical activity. Educational RAs have been deployed as museum exhibits [Bickmore et al. 2013] and as empathetic artificial tutors [Castellano et al. 2013].

Our work potentially relates to all the systems described, but none of them focused on the impact of the first user–agent interaction.

First Impressions. Maat et al. [2009, 2010] showed how different turn-taking strategies in a user–agent interaction could influence users' impressions of an agent's personality (agreeableness), emotion, and social attitudes. Effects of agent gaze behavior [Fukayama et al. 2002] and eye blinking rates [Takashima et al. 2008] on user's impressions of affiliation (friendliness, warmth), nervousness, and intelligence were also studied. Most recently, Bergmann et al. [2012] investigated how appearance and nonverbal immediacy affect the perceived warmth and competence of virtual agents over time.

In this work, we are dealing with the very first formative moments of an encounter, as opposed to the agent turn-taking strategies. And we are looking at a wider set of nonverbal immediacy behaviors rather than solely gaze and eye-blinking behavior. The work of Bergmann et al. complements our research, but they focused on gestures in a task-related agent monologue.

Expression of Personality and Interpersonal Attitudes. There are several personality perception studies focusing on expression of agent's extraversion [2010, 2014], neuroticism [Neff et al. 2011], and all Big 5 traits (except openness to experience) [Doce et al. 2010] with verbal and nonverbal behavior. A real-time back-channel selection algorithm for choosing the type and frequency of back-channels to be displayed according to the personality of an agent is described in Sevin et al. [2010].

Gillies et al. [2004] animated agents' nonverbal behavior in improvisational visual media production, expressing interpersonal attitudes toward one another. Lee and Marsella [2011] modeled by-standers body language and gaze to express their attitudes in an interactive mixed-reality drama scenario set in a wild west saloon. Finally, a computational model to express interpersonal attitude for a virtual coach by selecting sequences of nonverbal behavior (e.g., a head nod followed by a smile) has been proposed by Chollet et al. [2014].

These works incorporated personality traits or interpersonal attitudes separately into virtual agents mainly designed for face-to-face interactions or interactive drama. In this paper, we focus on user's judgments of nonverbal behavior when both personality (extraversion) and attitude (affiliation) are expressed at the same time. Furthermore, we studied the user's interpretation of the nonverbal behavior exhibited by an agent in the formative moments of their first virtual encounter.

Impact of User's Personality. The user's own personality can impact the interpretation of the nonverbal behavior exhibited by an agent. Studies in human–robot and human–agent interaction that targeted the benefits of a match between agent and user personality have yielded inconsistent results. Isbister and Nass [2000] (*static still figure*) found that people tended to prefer a virtual character figure whose personality was complementary to their own, whereas Nass et al. [1995] (*animated computer interface*) and Tapus et al. [2008] (*robot*) showed that people preferred to interact with personalities similar to their own. According to Andre and colleagues [2000], an agent should react to a negative user's affective state with a more extraverted and agreeable personality. Bickmore and Cassell [2001] showed that an agent using small talk when interacting with users increased trust in it for those extraverted, but for the introverted it had no effect.

Users' personality has influence on agent's evaluation [Pütten et al. 2010]. In particular, agreeable and extraverted users judged agents more positively compared to those who were shy. Kang and colleagues suggested that users' shyness [2008b] and Big 5 personality traits [2008a] are associated with their feelings of rapport when they interacted with different versions of virtual agents capable of exhibiting different styles of nonverbal listening behavior.

In this work, we are interested in the possible moderating effect that user's personality may have on judgments of agent's extraversion and affiliation. The novelty consists of studying the possible interplay between these two user judgments, whereas users observe and interpret an agent's nonverbal behavior even before the face-to-face interaction takes place.

User-Agent Interpersonal Distance. The RAs REA [Cassell et al. [1999, 2000]] and Tinker [Bickmore et al. 2013] both had life-sized articulated graphical bodies and acknowledged the user's presence using motion trackers. REA would first turn to face an approaching user and then attempt to salute them. Users could respond by coming closer and by speaking and gesturing, more or less choosing their preferred interaction distance. Tinker provided a touch screen for receiving user's input utterances. It used a motion sensor coupled with a hand biometric device for detecting surrounding users and (re)identify them using hand biometric features.

Ring et al. [2014a] developed an RA for isolated older adults. Only the agent face was displayed on a touch screen monitor that users used to input their utterances by choosing among several options. They developed a *passive* version, in which conversations with the agent had to be initiated by the user by touching an option on the touch screen. Then they build a *proactive* version, in which the agent could detect when users walked by the system via a motion sensor and attempted to initiate a conversation by verbally greeting them. Results from a pilot study indicated that the proactive version was more accepted compared to the passive one relying upon users to initiate interactions.

Deshmuk et al. [2010] worked on a robotic RA capable of engaging in spontaneous interaction with the users by approaching them in a proper manner (i.e., asses whether the interaction could be initiated even when not requested) or be ready to start an interaction if the user showed willingness to engage with the robot (only when the user was within its social area). Heenan et al. [2013] adopted Kendon's greeting model and Hall's proxemics theory to provide a humanoid robotic agent with social awareness toward approaching users. The physical constraints of the robot prevented it from realizing facial expressions and eye gaze behavior; however, proxemics, gaze (whole head), and simple gesturing (e.g. hand wave) were exhibited during the greeting process with the user.

Koay and colleagues [2014] studied users' proxemic preferences for a domestic non-humanoid service robot that handed them an object. They conducted a field study in a home setting and discovered that social expectations impacted the results even if a non-humanoid robot was deployed. In particular, findings indicated that participants were comfortable with the robot approaching to the closest implemented distance (0.5 m) and preferred the robot to approach them from the front. In a similar longitudinal study, they found that humans preferred to approach a robot more closely than they allowed the robot to approach them in a physically restricted area [Walters et al. 2011]. Moreover, most of the user–robot proxemics adaptation occurred in the first two interaction sessions, and for the remaining four weeks, approach distance preferences remained relatively steady.

Finally, Kim and colleagues [2014] studied proxemics behavior (close vs. distant) and social status (supervisor vs. subordinate) of an anthropomorphic robot in a card-matching game scenario wherein participants played with the robot. They found that robot's social status interacted with proxemics behavior on the overall judgments of users' experience. In particular, participants who interacted with the supervisor robot judged the experience more positive when it was close, whereas interactions with the subordinate robot resulted in a more positive experience when the robot was distant.

Independently from the agent's embodiment, these works suggest that considering interpersonal distance between user and agent and, in particular, paying attention to the nonverbal behavior exhibited by the agent during the initial moments of interaction (e.g., the greeting process) are important aspects to factor in the design of conversational interfaces. None of these systems exhibited different behaviors as a function of decreasing interpersonal distance, and none were used to study the effects of subtle nonverbal manipulation on users' first impressions.

Public Space Deployments. Virtual and robotic conversational agents have been deployed in public spaces (mainly museums) for field studies and for collecting large amounts of data.

Virtual Agents. Kopp et al. [2005] installed *Max* in the Heinz Nixdorf Museums Forum in Germany. Max was projected on a life-size screen, standing face-to-face with visitors. Its expressed behavior included a German voice accompanied by conversational gestures and gaze. In Robinson et al. [2008], they explore the issues of deploying conversational agents in public spaces through an analysis of a corpus of visitors–agent interactions collected at the Cooper Hewitt Museum in New York.

Lim et al. [2010] deployed *Sarah* on a large display in a meeting room of Heriot-Watt University in UK. This was an exploratory study "in the wild" for gathering information about interesting conversation topics in order to inform the design of an autonomous version of the agent that can act as a long-term companion in a natural social setting. Students did not interact enough time with Sarah in order to have a consistent amount of information. According to the authors, first impressions played an important role: "due to an initial disappointment, the students might find Sarah not worth interacting with... Moreover, the attention they gave to the agent and the time spent interacting with Sarah was too short to create a valid impression of her."

In Nunamaker et al. [2011], they present an automated conversational kiosk that uses a virtual agent to interview individuals and detect changes in arousal, behavior, and cognitive effort by using psycho-physiological information systems. *Ada and Grace* [Traum et al. 2012] were twin virtual conversational agents at the Boston MoS. They could interact with each other and were designed to engage visitors using unrestricted speech input from a microphone. A virtual 3D cartoon-style pedagogical agent was also exhibited at MoS [Lane et al. 2011]. It was designed to be approachable, supportive, and understanding in order to improve the experience of museum visitors in that area. Finally, Samer et al. [2012] deployed *Furhat* at the London Science Museum as part of a robot festival. Furhat is a 3D back-projected human like robot head that utilizes a computer animated face to carry multimodal multiparty interactions with visitors.

Robotic Agents. Gockley and colleagues [2005] presented *Valerie*, a robot receptionist installed in the entrance way to Newell–Simon Hall at Carnegie Mellon University, in USA. Valerie was able to give directions to visitors and look-up the weather forecast while also exhibiting a compelling personality and character to encourage multiple visits over extended periods of time. Ludewing and colleagues conducted a field study using a shopping robot with extraverted personality [Ludewig et al. 2012]. They found that the extraverted robot received a higher social acceptance by the users than a conventional version (i.e., close to neutrality).

These public deployments had virtual and robotic embodiment capable of multimodal interaction. However, none of them incorporated a model of first impressions

Fig. 2. The setting of our study with the user's avatar entering the virtual museum entrance in first person mode and the greeting agent waiting inside. The schematic shows points where specific behaviors were exhibited by the agent during the avatar's approach.

through nonverbal behavior when a user approached them for the first time during the initiation of the interaction (i.e., greeting process).

5. PREPARATORY USER STUDIES

5.1. Nonverbal Behavior Interpretation Study

In this first preparatory study our subjects, represented by an on-screen avatar, approached a series of greeting agents in a 3D virtual museum entrance displayed on a 19" LCD monitor (see Figure 2). The goal was to evaluate users' impressions of a greeting agent's extraversion and affiliation in a first encounter. The agents exclusively exhibited a set of nonverbal immediacy cues that were systematically manipulated during the initiation of the interaction. The approaches had a duration of 12.5s. This was chosen after a prior validation study described in the electronic appendix.

The main research questions were the following.

(1) How do users interpret an agent's smile, gaze, and proxemics behavior in terms of first impressions of extraversion and affiliation?
(2) Does the user's own personality influence the interpretation?

We had two additional exploratory questions: (1) In addition to personality and interpersonal attitudes, what kind of judgments are people forming when observing and interacting with conversational agents that display certain nonverbal cues? (2) When moving from a first to third person camera perspective (user's avatar visible from the back), do users interpret observed nonverbal behavior in the same way? For this second question, the study was split into two trials, 1P and 3P, differing only in the camera perspective used.

5.1.1. Hypotheses and Setting. The following are the two hypotheses that we had for both trials:

—H1: The amount of *extraversion* that subjects attribute to a greeting agent (a) depends on the displayed smile and proxemics behavior and the amount of gaze it exhibits toward the subject during the initiation of the interaction and (b) is further moderated by the subject's own personality.

—H2: The amount of *friendliness* that subjects attribute to a greeting agent (a) depends on the displayed smile and proxemics behavior and the amount of gaze it exhibits toward the subject during the initiation of the interaction and (b) is further moderated by the subject's own personality.

The context of interaction was the main entrance of a virtual museum. The scene always started with the subject's avatar (AVATAR) outside, in front of automatic sliding doors, and the greeting agent (AGENT) standing inside, close to a reception desk watching a computer screen. Figure 2 (on the left) shows this setting in first person perspective when the approach has already started.

We designed our stimuli by keeping in mind the subtlety of the agent's exhibited behavior. This study is described in detail in Cafaro et al. [2012], but we remark here how our three independent variables differed. These variables were *smile* (no vs. yes), *gaze* (low % vs. high %), and *proxemics* (no step vs. step).

Figure 2 (on the right) shows a schematic top view of the scene with the AVATAR and the AGENT in their initial positions. The black arcs are points where specific behaviors were exhibited according to Kendon's greeting model and Hall's proxemics theory, distances (in meters using the American notation) are from the AGENT to the arch.

We created a *baseline behavior* for the AGENT that was exhibited across all conditions of the study when the AVATAR approached it. This consisted of watching a computer screen at the beginning with both head and eyes toward it, gazing at the AVATAR for 2s when it was at T1 (8m), looking back at the screen moving only the eyes and, finally, gazing at the AVATAR at T3 (3.30m). The AVATAR always stopped at T4 (1.43m).

In the smiling condition, the AGENT exhibited a polite smile starting at T1, when not smiling a neutral facial expression was displayed. The smiling facial expressions was created by controlling Facial Action Units [Ekman and Friesen 1978] applied to the neutral facial expression and affecting the agent's lips, cheeks, and eyebrows. The *"high %"* gaze was obtained with a 2s eye glance at T2 (on top of the baseline) that was not present in the *"low %"* gaze condition. It follows that the difference between "low %" and "high %" gaze conditions was simply related to their duration, in the former the AGENT looked at the subject's AVATAR for a shorter time compared to the latter, although it was never staring at the user for the whole interaction. The *"step"* condition was simply a step toward the AVATAR when it was at T3 keeping the arms behind the back.

5.1.2. Main Findings. We had 32 participants in each trial (1P: 20 males and 12 females; 3P: 19 males and 13 females). The detailed statistical analysis can be found in Cafaro et al. [2012], we review here the main findings. This study has provided evidence that people form impressions of a greeting agent's extraversion and friendliness in the very first moments of a virtual encounter when the agent exclusively exhibited nonverbal immediacy cues (smile, gaze, and proxemics). The fact that main effects were found suggests the power of even subtle behavioral variations.

We also discovered that the specific cues adopted had well-defined and separate meanings along the two dimensions measured, in particular, proxemics behavior had the biggest impact on impressions of extraversion, whereas smile and gaze had significant impact on friendliness. Furthermore, we demonstrated that camera perspective (first or third view) has no influence on user interpretations of an agent's nonverbal behavior in this context.

Influence of user's personality on behavior interpretation was found. Our results showed influences of subject's extraversion and agreeableness traits on behavior interpretation, except for the neuroticism trait. It is possible that user's preferences for a virtual guide agent might depend on those personality traits. Considering that the social dimensions of our model (i.e., extraversion and friendliness) are likely to relate to these traits, we kept investigating those in the second laboratory study and excluded the neuroticism trait.

We further looked at possible subjects' gender and cultural identity effects in order to explain possible influences of such factors in addition to subject's personality on the behavior interpretation. The cultural identity of the subjects was obtained, as part of the demographic information, by asking participants to select the country that most represented their cultural identity from a list of all countries in the world. Although labeling someone's culture with a country is not representative of many facets of a given culture (e.g., consider sub-cultural identities depending on local/regional dialects), this single choice (i.e., nationality label) allowed us to obtain the information in a simple but not naive (i.e., 1-to-1 mapping between nationality and culture) manner, by giving subjects the freedom to choose their cultural background while still keeping it simple and non-ambiguous simply by picking a label that best represented their cultural identity (i.e., the closest one).

We didn't find significant interaction effects to report for the subjects' gender analysis in either trial (detailed statistics for this further analysis are provided in the electronic appendix). For the cultural identity, we had two sub-groups reflecting our largest group of subjects (i.e., *Icelanders*) and the remaining participants (i.e., *Others*). The analyses revealed a difference in the smile perceptions between the two groups. *Icelanders* interpreted smiling agents as more introverted compared to those not smiling, whereas the *Others* group interpreted smiling agents as more extraverted than non-smiling ones. In sum, cultural differences did occur, but with a small effect size compared to the main effects of our stimuli (i.e., smile, gaze, and proxemics behavior).

The findings of this first study indicate that results in social psychology research on the assessment of personality traits and attitudes on the basis of nonverbal behavior [Burgoon et al. 1984; Riggio and Friedman 1986; Argyle 1988] do translate to the context of user–agent interaction. In particular, outcomes of using nonverbal immediacy [Richmond et al. 2008] are preserved in virtual encounters.

Finally, *"Agent Likeability"* results mirrored those of *"Agent Friendliness"*, smiling and gazing agents were considered more approachable and likable. According to our measurement of likability, these results predict that subjects might consider friendliness (attitude) more important than extraversion (personality) when it comes to deciding whether to continue the interaction or not later.

5.1.3. Impact of Results on First Impressions Model. We can conclude that when exhibiting smile, gaze, and proxemics behavior in orchestration, we can manage an agent's given impressions of friendliness and extraversion simultaneously, by focusing on smile and gaze for the former and proxemics for the latter. It is remarkable that these impressions can be managed at the same time, therefore without risking to show off an unwanted impression (e.g. unfriendliness) when attempting to be, for example, more reserved by waiting for the user to come close without stepping toward him/her, thus not being judged as extroverted if required by the social context.

In our first impressions model of personality and attitude, we now know how agent's nonverbal greeting cues are interpreted, but the impact and relative importance of those impressions on user's acceptance remain undiscovered. The study presented next addresses these aspects building on these results but moves toward a setting closer to the final field study where a life-sized virtual agent is displayed.

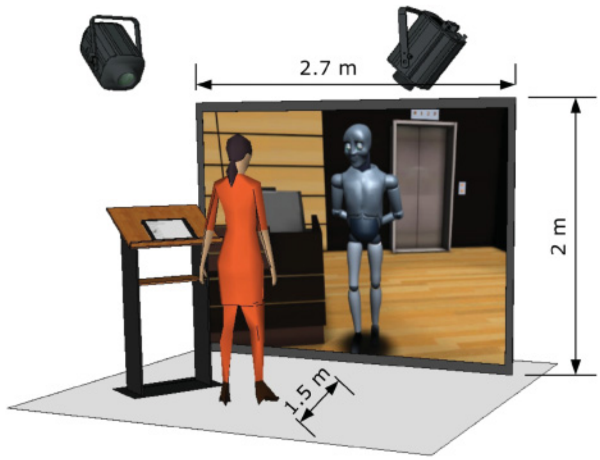

Fig. 3. A 3D reconstruction of the setting adopted for our study showing the user position standing still in front of an animated view of the full-size guides. The tablet computer on the kiosk at the left of the user was used to answer the questions after meeting each guide in the virtual museum entrance shown on the screen in front of the user.

5.2. Nonverbal Behavior Impact Study

This second study aimed at investigating the impact of personality and attitude impressions of an agent on users' relational decisions in terms of how likely they are to interact with the agent again and how often they would then interact with it. We also wanted to study possible effects of concordance between the user's own personality and the agent's personality and attitude when making these decisions, i.e., whether the user's relational decisions depend on his/her own personality.

In order to make the presentation of our stimuli clearer and, most importantly, make the setting more similar to the field study described next, subjects observed a series of animated views of first-person perspective approaches to life-sized agents presented as guides of a virtual museum. The guides exhibited two levels of personality (high vs. low extraversion) and attitude (high vs. low friendliness) toward the subjects during initial greeting approaches of 12.5s. The manipulations were exclusively based on nonverbal immediacy cues exhibited by the guides as suggested by our previous study. After meeting each guide, subjects were immediately asked to express, in general, how likely they were to spend time with it again and their preferences about the number of guided tours they would be willing to take with the agent. The study was first summarized in Cafaro et al. [2013], but more detail is provided in the following sections.

5.2.1. Apparatus and Stimuli. The subjects interacted with our guides standing at a fixed location in front of a 2.7m wide and 2.0m high back-projected screen at a distance of 1.50m. They experienced the stimuli without using stereoscopy. Figure 3 shows a 3D representation of this setting. The screen projected life-size agents in a 3D virtual entrance of a museum. The scene always started with the subject outside, in front of automatic sliding doors and a single guide agent waiting inside, standing close to a reception desk, holding its arms behind its back and watching a computer screen, similar to our previous study.

From the subject's point of view, an experimental session consisted of observing (in first person perspective) their approach to a series of virtual guides to symbolically give each of them a business card, as shown in Figure 3.

A touch screen tablet computer was used to give subjects the ability to start each interaction, by pressing a button on the screen, and to administer self-report

Table I. How Independent Variables (IVs) Map to Nonverbal Cues Based on Our Previous Findings

Guide level of		
Extraversion	Friendliness	Nonverbal cues
Low	Low	No smile, low % gaze, no step
Low	High	Smile, high % gaze, no step
High	Low	No smile, low % gaze, step
High	High	Smile, high % gaze, step

Table II. Summary of Second Study Measures. Likelihood of Encounters is on a Five-Point Scale

Measure	Question	Range
Likelihood of encounters	Would you like to do business with this guide?	1 (No, definitely not) 5 (Yes, definitely)
Number of visits	How many guided visits would you like to take with this guide?	0–10
Guide preference	Of the four guides you have just met, which one would you prefer as your guide?	Guide [1-4] or "I don't know"
Subject personality	Extraversion and agreeableness traits	

questionnaires at the end of each interaction. When a subject touched the start button on the tablet, it triggered an animated locomotion toward the guide agent. It automatically stopped when the subject reached the encounter space of the agent. The self-report questions were displayed after approaching each guide (with the agent no longer present on the screen) and subjects tapped on the tablet screen to reply.

A system of lights in the ceiling of the experiment room was controlled via software to lead the subject through the various steps of the study. A light would turn on in front of the screen to make a marker visible on the floor. This marker indicated the point where the subject had to stand when observing the approach toward the guide on the screen. Another light would turn on (switching the other off) to show that questions were available on the tablet right after approaching each guide. Pictures of the experimental room showing this setting are provided in the electronic appendix.

The independent variables were *Guide Extraversion* (low vs. high) and *Guide Friendliness* (low vs. high). The different levels of extraversion and friendliness were obtained by manipulating agents' nonverbal immediacy cues of smile, gaze, and proxemics according to our previous results. Since we discovered that agents that stepped toward approaching users were judged as more extraverted, our guides did the same when *guide extraversion* was at *high* level and did not step toward the subjects when at *low*. On the other hand, we earlier found that smiling and gazing agents were judged as more friendly; therefore, we had smiling and more gazing (though with a few glances away as in the earlier study) guides for the *high* level of *guide friendliness*. The *low* level was obtained with the neutral facial expression and a lower amount of gaze at the user. Table I shows all the mappings between the levels of our IVs and the resulting nonverbal behavior exhibited by the guides. To control for ordering and carryover effects, we showed the resulting four different conditions with a full counterbalanced treatment order [Bordens and Abbott 2002] in a within-subjects design.

5.2.2. Measures. Table II provides a summary of our measures. The first two measures assessed relational decisions: *Likelihood of Encounters* assessed the overall willingness to visit the agent again, whereas *Number of Visits* assessed decisions about the frequency of those visits. After meeting the four guides, we asked subjects to choose their preferred guide (*Guide Preference*) among the four that they met and we assessed their personality (*Subject Personality*) using Saucier's Mini-Markers [Saucier 1994] set of adjectives for extraversion and agreeableness traits. The full set of questionnaire forms administered is listed in the electronic appendix.

5.2.3. Participants and Procedure. We had 24 participants recruited via public announcements in our university campus and the surrounding city. There were 15 males and 9 females. 62% chose *Iceland* as cultural identity. They were aged 21–60 with 54% in the 21–30 range. All subjects were well educated and most were at least familiar with computer science and psychology. Detailed demographic data are provided in the electronic appendix. They were led to a dedicated room at our university facility, instructed about the procedure, shown a tutorial for familiarization, and asked to sign the consent declaration. Subjects were led to believe that the selected number of future visits represented an actual time commitment. The following is an excerpt of the consent form that we used.

> By signing this document, I agree to come back in these facilities and be guided in one or more virtual tours of the museum according to the particular guide that will be assigned to me and the preference (number of visits) that I expressed for him. The guided tours (if any) will be scheduled over a period of two consecutive months at my earliest convenience. The start date of this period and all appointments will be scheduled in concordance with the investigator. Every visit will require approximately 15 minutes and different area of the museum will be shown if more than one visit is scheduled.

We explained to our subjects that the assignment of the guide would be random, but we would respect their wishes regarding each one (i.e., how often they would like to see them). This was introduced to prevent subjects expressing realistic preferences for only the most preferred guide and giving zero or a low number of visits to the others. After signing the consent declaration, we clarified again that a tour would require them to come back to our facilities and spend time with the guide delivering it. Then the investigator left the room to monitor the session from an adjacent observation room. The session consisted of (1) meeting each guide and then replying to questions on the tablet including the likelihood of future encounters and the number of visits, (2) choosing the preferred guide, (3) completing the personality inventory, and (4) providing demographic data. Finally, the investigator debriefed them, revealing the nature of the deception and reaffirming consent for using the data. For more details about the documents used for this study, see the electronic appendix.

The first five subjects of our study were quickly interviewed after the debriefing. Our goal was to check whether (a) they believed in the real-time commitment that they were taking by giving different options for the number of visits, (b) the choice was not influenced by their future availability, and (c) the virtual museum domain and the possibility of having guide tours were of sufficient interest. Only one subject reported having limited time availability in the near future, but all the others stated that their choices were only influenced by the guides presented. They all claimed that they seriously considered the possibility of coming back for subsequent visits and that the museum scenario had relatively little effect on their decisions compared to the guides' behavior.

5.2.4. Hypotheses. From the results of our previous study and findings in social psychology on the outcome of nonverbal immediacy [Richmond et al. 2008; Burgoon et al. 1984], we predicted that (a) friendliness, as an interpersonal attitude and a "short-term" feature, would have a main effect for all our measurements, whereas (b) extraversion, as a personality trait of the guides and a "long-term" feature, would interact with the subject's own personality. In the absence of consistent evidence on user–agent personality matching/mismatching preferences (as discussed in Section 4), and in accordance with a recent study in the conversational agents domain Cerekovic et al. [2014], we opted for the matching school of thought. Thus, we predicted that similarity in the agent's personality (*Guide Extraversion*) and the user's personality would have positive effects on outcomes. These were our hypotheses.

—*H1 (Likelihood of Encounters)*: (a) Guide Friendliness will have a main effect on the likelihood of future encounters. Subjects will be more likely to encounter *high* friendliness guides again compared to the *low* friendliness ones; (b) subjects own personality, either the extraversion or the agreeableness trait, will positively interact with guide extraversion (e.g., *High* extraverted subjects will be more likely to encounter *high* extraverted guides again);
—*H2 (Number of Visits)*: (a) Guide Friendliness will have a main effect on the number of future visits. Subjects will choose a higher number of visits with *high* friendliness guides compared to the *low* friendliness ones; (b) subjects own personality, either the extraversion or the agreeableness trait, will positively interact with guide extraversion (e.g., *High* extraverted subjects will choose a higher number of visits with *high* extraverted guides);
—*H3 (Guide Preference)*: The preferred guide will have *high* level of friendliness.

5.2.5. Results. We conducted separate mixed-design analysis of variance (ANOVAs) for the two dependent variables *Likelihood of Encounters* and *Number of Visits*, with guide extraversion and guide friendliness as within-subjects factors and subject extraversion and agreeableness as between-subjects factors. We used a factorial model, but we omitted interactions among the between-subject factors. In order to use the two subject personality traits as between-subjects factors, for each measured trait we split our population at the median into *"low"* and *"high"* groups.

Likelihood of Encounters. A mixed-design ANOVA was conducted for this measure. The analysis revealed a significant main effect of guide friendliness, $F(1, 21) = 21.91$, $p = 0.000$, $\eta_p^2 = 0.51$; subjects were more likely to do business with *high* friendliness guides later, compared to the *low* friendliness ones (*H1-a supported*). The main effect of guide extraversion was not significant ($F(1, 21) = 0.39$, $p = 0.54$, $\eta_p^2 = 0.018$). No significant interaction effects were observed between the guide extraversion and subject personality traits; therefore, *H1-b* was *rejected*.

We also looked at possible subjects' gender and cultural identity effects conducting an additional analysis with these two as between-subjects factors (as also done in the previous study). However, we did not find significant interaction effects to report.

Number of Visits. This variable was measured on an ordinal scale ranging from 0 to 10 and did not follow a normal distribution. In order to use standard analysis techniques for repeated measures (i.e., RM-ANOVA), we applied an aligned rank transform (ART) for non-parametric factorial data analysis as suggested by Wobbrock et al. [2011]. The analysis of the transformed values revealed a significant main effect of guide friendliness on the number of visits, $F(1, 20) = 14.22$, $p \leq 0.001$, $\eta_p^2 = 0.47$; subjects preferred to take a higher number of visits with *high* friendliness guides compared to the *low* friendliness ones (*H2-a supported*). The main effect of guide extraversion was not significant ($F(1, 20) = 0.062$, $p = 0.80$, $\eta_p^2 = 0.003$). No significant interaction effects between Guide Extraversion and the subject personality traits were observed; therefore *H2-b* was *rejected*. There weren't significant interaction effects on the number of visits between Guide Extraversion (Friendliness) and the two factors subjects' gender and cultural identity.

Table III. The Number of Preferences (PREFS) Received by Each of the Four Types of Guides Presented in Our Study

	Guide level of		
Type	Extraversion	Friendliness	Prefs
1	Low	Low	3
2	Low	High	10
3	High	Low	0
4	High	High	10
5	"I don't know"		1

Guide Preference. All the subjects, except one, had a preference for a specific guide type among the four presented as shown in Table III. Subjects showed a highly significant preference for the guides with a *high* level of friendliness, $\chi^2(2, N = 23) = 12.56$, $p < 0.001$; therefore *H3* is *supported*. For the chi-square test of goodness-of-fit that was performed we dropped the single subject without preference and grouped the other preferences in two categories depending on the guide's level of friendliness.

Frequency Analysis for Number of Visits. In support of earlier results and to examine this measure from a different angle, we treated the number of visits' preferences received by each guide as ordered categories and computed the frequencies of choices for every single category. For high friendliness guides—those most often preferred—subjects more frequently selected a high number of visits (5–8), compared to low friendliness guides where a low number of visits (0–1) were more commonly selected. The modes for the high friendliness guides were three visits (at *low* extraversion) and two visits (at *high* extraversion): A relatively high number of subjects would have spent from 30 (2×15) up to 45 (3×15)min with them later, considering that each visit would have required 15min.

5.2.6. Discussion. The friendliness of the virtual guide had a main effect for all our measurements, thus supporting our hypotheses (H1-a, H2-a, and H3). In particular, *high* friendliness received a higher number of visit preferences than *low* friendliness guides, subjects were more inclined to do business with them later, and they were the most preferred ones regardless of the level of personality that was exhibited.

This would suggest that first impressions and relational assessments in human-agent interactions are made quickly, often in the beginning moments of initial conversations or, in our case, even prior to a conversation as in human–human interaction [Sunnafrank and Ramirez 2004].

Contrary to what we expected, we did not observe any interaction effect between the user's own personality and the guide's level of extraversion, when measuring the number of visits and the likelihood of encounters (H1-b and H2-b rejected). A possible explanation is that considering the short amount of time that subjects had to form impressions of the guides, they based their decisions on a feature, interpersonal attitude, that happens in a "short term" and is specifically addressed to them in that particular context, thus fulfilling their expectations about the guides being approachable, friendly and polite.

Furthermore, there is a very little difference between the affiliation dimension of interpersonal attitudes and the agreeableness personality trait in terms of expressiveness through nonverbal behavior only [Richmond et al. 2008; Argyle 1988]. There is evidence that agreeableness might be related to formation of alliances and cooperative relationships [Mehu et al. 2008]. Thus, we could have expected to have such interaction effects by focusing on manipulating the guide agreeableness level, instead of extraversion.

One limitation that should be considered is the sample size (24 subjects) needed for an adequate analysis of concordance between the user's own personality and the agent's personality and attitude. Even though results in the earlier experiment were promising, a bigger sample might have revealed more consistent results. We searched for other factors that might have possibly had an influence, such as subjects' gender and cultural identity, but no interactions were found. We emphasize, however, that for a within-subjects design our sample size sufficed, in fact yielded large effect sizes for the main two measurements, i.e., Likelihood of Encounters ($\eta_p^2 = 0.51$) and Number of Visits ($\eta_p^2 = 0.47$).

Another limitation concerns our number of visits outcome. It was a self-report measure and, even though our subjects consented to return to our lab according to the preferences they expressed, it was still a hypothetical decision. A behavioral measure, for example assessing whether subjects actually return for the tours they agreed to, would be preferable. This might be done within a longitudinal study design that would also allow us to examine the stability and long-term impact of attitudes and impressions formed in these brief user-agent encounters.

5.2.7. Main Findings. Our major finding is that the impressions of our virtual agents, exhibiting the proper nonverbal immediacy cues at specific points during only 12.5s of a first greeting encounter, had significant effects on users' relational decisions in terms of choosing how likely and for how often they would spend time with the guides later. In particular, guide friendliness, expressed with smiling and gazing more at the subjects, had a main effect for all our measures.

Furthermore, it is remarkable that all subjects, except for one, upon agreeing to come back in the consent, were able to choose their preferred guide among the four presented even though the interaction time was short, they had never seen our virtual guides before, and they also had the option to not choose any of them.

5.2.8. Impact of Results on First Impressions Model. From this second preparatory study, we learned that in a virtual greeting encounter, the user's impressions of a guide's interpersonal attitude have greater impact on the likelihood and frequency of subsequent encounter compared to impressions of a guide's personality. From the behavior interpretation study, we discovered that smiling and gazing behavior suffice for managing those impressions (i.e. friendliness) when initiating the interaction.

6. MAINTAINING ENGAGEMENT AND MANAGING FIRST IMPRESSIONS IN A PUBLIC SPACE

Our ultimate goal in this work was to test our model of first impression management in a real-world setting, in which first impressions can have a significant impact on task outcomes. Museums represent an ideal context for this evaluation. Despite the use of visually appealing, interactive installations, guests in modern museums spend only 3–5min at un-staffed exhibits, making it difficult to achieve significant levels of in-depth education about any given topic. To address this problem, we designed a relational museum guide agent—named "Tinker"—that uses human social and relational behaviors to both draw guests to the exhibit and keep them there as long as possible so that learning can occur. This museum agent incorporates both our first impression model—to draw users to the exhibit—and theories of relationship construction, to keep users at the exhibit as long as possible.

In the following sections, we first provide a description of the Tinker exhibit in the chosen public space (Section 6.1). Then, we summarize results of a former field study on maintaining engagement with relational behavior once users have decided to interact with the system (Section 6.2). Finally, in Section 6.3, we present the field study

Fig. 4. Tinker character appears as a robotic agent. The large scrolling text screen placed behind Tinker showed the content of the last several conversational turns, the smaller sign on the right displayed system status information and a demonstration animation sequence showing approaching visitors how to use the hand reader.

with Tinker where we investigate whether the use of our first impression model can influence initial user attraction to the system (i.e., encourage engaging in interaction).

6.1. The Tinker Exhibit at the Boston Museum of Science

Tinker is a virtual museum guide that was installed in the Computer Place exhibit at the Boston MoS[1], in Massachusetts, USA. It was able to provide visitors with information on and directions to a range of exhibits in the museum as well as discuss the theory and implementation underlying its own creation. The exhibit was operational from September 2007 to May 2014 and has been seen by over 170,000 museum visitors [Bickmore et al. 2013]. Our field study ran approximately for 1 year (309 days) from December 2012 to October 2013.

The animated agent was displayed a 46” LCD display and it appeared in the form of a human-sized anthropomorphic robot that speaks using a synthetic voice (Figure 4). Tinker uses a biometric identification system (based on hand image features) so that it can re-identify visitors it has already talked to. It also uses a touch screen for user input, including multiple-choice selection of conversational utterances, user given name input, and topic navigation menus using iconic representations. A smaller sign behind Tinker displays system status information as well as a demonstration animation sequence showing approaching visitors how to use the hand reader. Information about a given visitor—including facts learned about them through dialogue, exhibits they asked about or said they were visiting, their computer literacy level, and the relationship model—is stored in a persistent data structure during an interaction, and saved to a database at the end of the interaction. When a return visitor is identified, the data for that visitor are reloaded into memory so the conversation, and the relationship, can be continued. More details on Tinker’s design can be found in Bickmore et al. [2013].

6.2. Study on Maintaining Engagement with Relational Behavior

An initial study was conducted (previously reported in Bickmore et al. [2013]) to determine whether simulated relationship-building behavior by Tinker would lead to increased retention at the exhibit, once users had already decided to visit it. A randomized, between-subjects experiment was conducted in which the following relational behaviors were either turned on or off:

Empathy. Empathy is one of the core processes in building and maintaining relationships [Havens 1986; Gelso and Hayes 1998]. Tinker expresses empathy for feeling

[1]See the web page at: http://www.mos.org.

states expressed or implied by a visitor. For example, after asking about a visitor's experience at the museum, a user response expressing boredom results in "I am sorry to hear that. I hope you can find some part of the museum that interests you" (with a concerned facial display).

Getting Acquainted. Tinker asks visitors about themselves, such as who they are visiting with, and where they are from Svennevig [1999].

Self-Disclosure and Reference to Common Ground. Tinker makes references to information disclosed by a visitor at subsequent points in the dialogue, to remind them of their shared knowledge and interaction history [Altman and Taylor 1973].

Reference to Shared Values and Beliefs. Tinker agreed with many of a visitor's expressed likes and dislikes [Gill et al. 1999]. For example, if the visitor indicates they are a Red Sox (Boston baseball team) fan, Tinker would say it is a fan as well (if the visitor does not indicate this, Tinker did not talk any further about the team).

Humor. Tinker interjects humor at appropriate points in the conversation [Cole and Bradac 1996].

Form of Address. Once visitors have entered their given name, Tinker uses it to greet them [Laver 1981]. It also greets them by name on return visits, if the biometric hand reader recognized them [Schulman et al. 2008] (e.g., "Hi Bob, welcome back!").

Expressing Liking of the User and the Interaction and Desire to Continue. During farewells and subsequent greetings, Tinker expresses liking of the user [Okun and Kantrowitz 2014].

6.2.1. Main Findings. The automated study was completed with 1,607 adult visitors over two years, with 63% randomized to the relational condition. Results indicated the relational version of Tinker resulted in significant improvements on all measures of engagement, compared to the non-relational version, including:

Time Spent at Exhibit. Visitors in the relational group spent significantly longer time at the exhibit, 5.76 vs. 4.95min, $t(1605) = 4.41$, $p < 0.001$, and returned more frequently, 1.13 vs. 1.09 sessions, $t(1605) = 2.65$, $p = 0.008$, compared to the non-relational group.

Self-Report. Responses on point-point scale measures filled out after interacting with Tinker indicated that visitors in the relational group were significantly more satisfied, $t(1605) = 5.06$, $p < 0.001$, and expressed a significantly greater desire to continue their interaction, $t(1605) = 5.22$, $p < 0.001$, compared to those in the non-relational group.

Learning. Visitors in the relational group felt they learned significantly more compared to those in the non-relational group, $t(1603) = 4.55$, $p < 0.001$, even though the educational content was the same in both conditions. Most importantly, visitors actually learned more from the relational Tinker, scoring significantly higher on a post-session knowledge test, $t(1602) = 2.35$, $p < 0.05$, compared to participants who interacted with the non-relational Tinker.

This study demonstrated that proper use of social behaviors by a system—in this case relational behaviors used by an anthropomorphic museum exhibit—can have a significant impact on user engagement, once users have decided to interact with the system. In the next study, we investigate whether the use of first impression management behaviors can influence initial user attraction to a system, using the same museum platform.

6.3. Study on First Impressions Management in the Museum

To evaluate the impact of our first impressions model on initial engagement, we extended Tinker with the ability to exhibit friendly greeting behaviors and conducted

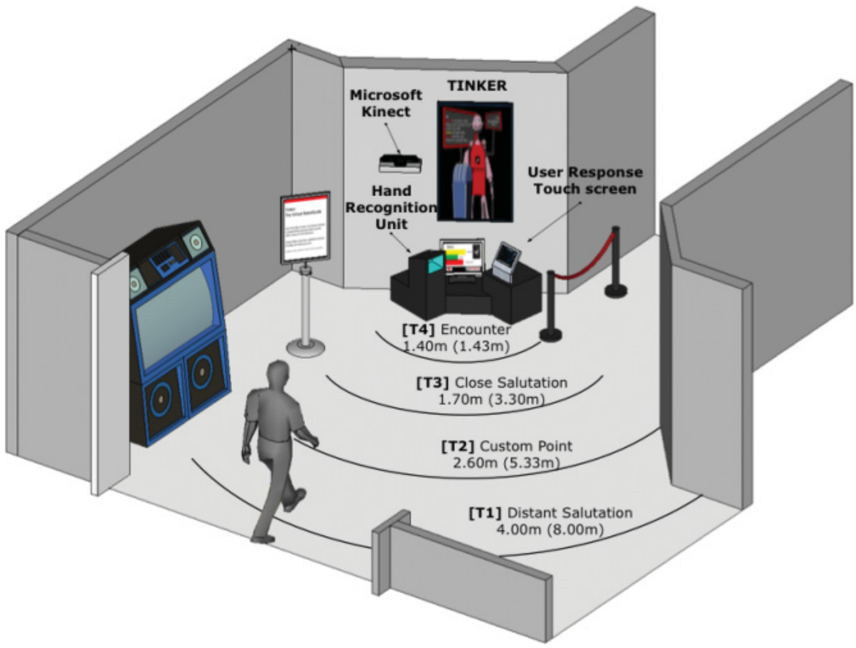

Fig. 5. A 3D reconstruction of the Computer Place exhibit's entrance at the Boston MoS showing Tinker's setup and points as suggested by Kendon's model where specific behaviors were exhibited by the agent during the visitor's approach.

a study to determine whether these had an impact on visitor decisions to approach and initiate interaction with the exhibit. We conducted a fully randomized experiment with a three-group, between-subjects design, in which we compared FRIENDLY and UNFRIENDLY versions of Tinker, that exhibited smiling and gaze cues according to our earlier studies, along with a CONTROL version of Tinker that did not exhibit any nonverbal reaction toward an approaching visitor (i.e., our model turned off).

6.3.1. Extensions to Tinker for First Impression Management. Applying our greeting model to Tinker, where users were free to move around in front of the exhibit, required both robust presence detection, to initiate a greeting or know when abandoned, and location tracking, to react according to the different stages in the greeting model. Figure 5 shows a 3D reconstruction of the area of the museum where Tinker was exhibited.

Due to space limitations in the museum and sensor limitations, we had to modify several parameters of our greeting model. We needed to scale down the distances suggested by Kendon by *50%*. This placed the *"Distant Salutation"* at a distance of 4m from Tinker (instead of 8m), which was the best compromise between the least distance where this reaction could have happened and the constraints due to the spatial arrangement of Tinker in the room. As we can see in picture 5, Tinker stood in a corner facing the main entrance of the Computer Place area (down left) and a corridor that leads to different exhibits (down right), both at distance of about 4m. We kept the original distance T4, instead of scaling it down to 0.78m, due to hardware sensing limitations.[2] Furthermore, the position in which visitors had access to the input devices (hand reader and touch screen) was at about 1.40m as well. Pictures of the Computer Place space where Tinker was exhibited are provided in the electronic appendix.

[2]We used a Microsoft Kinect and it stopped sensing users that were less than 1 meter away.

Table IV. The Three Groups in Our Between-subjects Study and the Corresponding Nonverbal Cues Exhibited by Tinker

Experimental conditions	
Group	Mapped nonverbal cues
Control	No reaction, Tinker wakes up from sleep
Unfriendly Tinker	Neutral face, low % of gaze at visitor
Friendly Tinker	Smiling face, high % of gaze at visitor

6.3.2. Experimental Conditions. The black arcs in Figure 5 are points, according to our model, where Tinker's nonverbal reactions were triggered during the visitor's approach as implemented in the preparatory studies. Both the new distances and the original model distances (in parentheses using American notation) are shown.

The nonverbal immediacy cues exhibited by the agent were smiling and gazing behavior. According to our previous findings, smiling and gazing more at the user yield impressions of high affiliation (i.e., a friendly attitude), whereas not smiling and a lower amount of eye contact were judged as low affiliation cues (i.e., an unfriendly attitude). Table IV summarizes the three conditions in this study and the corresponding nonverbal greeting behavior exhibited by Tinker.

In all conditions, Tinker beckoned the visitors when they reached the *"Encounter"* point T4 (1.4m) by demonstrating how to use the hand reader and inviting them to begin the interaction. We had a CONTROL GROUP with Tinker not reacting at all during the approach. In this condition, Tinker appeared to be asleep with the arms crossed and the head a bit tilted on the right side. In both FRIENDLY and UNFRIENDLY TINKER conditions, the agent woke up at T1 with a gaze at the visitors of 0.5s. Then it looked away (on the right side) and looked back at the visitors when they were at T3 (1.7m). In the FRIENDLY TINKER condition, we added a smiling facial expression at T1 (4m) and the *"high %"* of gaze behavior was obtained with a 0.5s eye glance at T2 (2.6m). The gazes at the visitors had a shorter duration (0.5s) compared to the original one (2s) used in the previous studies. This allowed Tinker to properly exhibit the gazes at each visitor given that we scaled down the distances to trigger the behavior. The second glance at the visitors at T2 in the FRIENDLY TINKER condition served to indicate continuous attention to the visitors.

Tinker's nonverbal behaviors were all displayed using pre-made animations. We conducted an informal manipulation check in our lab facilities ($N = 10$, 3 females and 7 males) with a deployment of Tinker similar to the exhibit at the museum, except for the new proximity sensing capability. We asked subjects to approach Tinker while we were testing each behavior (i.e., smile and gaze) separately on both levels (neutral vs. smiling facial expression and low % vs. high % of gaze) to verify that differences between the levels were correctly perceived by subjects (see the electronic appendix for details).

6.3.3. Measures. A summary of our measures is provided in Table V. As Kendon suggests, how far one goes out of one's way to meet another appears to have a communicational significance [Kendon 1990] and, most importantly, manifests the desire to engage in an interaction.

Visitor Actions. This outcome measured how far out of their way visitors went when approaching Tinker. It is divided into four possible outcomes that can be sequentially performed by a visitor and eventually lead to a conversation session with the agent. Visitors quitting (i.e., walking away) *at any point* during the approach performed no actions (i.e., *zero*). Those completing the approach, thus arriving at the *"Encounter"* point T4, were performing one action. At that point Tinker invited the visitor to begin the session by placing the hand in the reader, thus the second action was performed

Table V. Summary of Measures. The Actions Taken by an Approaching Visitor Leading to a Conversation with Tinker are Divided up into Four Possible Outcomes. Visitors Quitting while Approaching (i.e. Walking Away) Performed Zero Actions. Once the Approach was Completed at T4 (One Action), the Visitor Could Place the Hand in the Reader (Two Actions) and After Listening to Tinker's Instructions the Conversation Began (Three Actions). The Second Outcome was Conditional to the Number of Actions Accomplished Earlier and Measured the Duration of the Conversation that Followed

Measure	Description	Range
Visitor actions	Actions completed by the visitor before engaging in a conversation session with Tinker: (0) Visitor walks away; (1) Approach completed at T4; (2) Hand in the reader; (3) Conversation session can start.	[0–3]
Session duration	Duration (in seconds) of a visitor conversation session with Tinker	[0–∞]

when the visitor did that. Finally, after listening to Tinker's instructions on how to keep the hand in the reader and choose dialogue options from the touch screen input, the visitor could begin the conversation session with the agent. The moment the visitor began the session signaled the last possible action that we measured.

Session Duration. This is a conditional outcome dependent on the number of actions taken before. It measures the duration of conversation sessions in seconds; however; those are measured only for visitors that completed all of the actions prior to beginning a conversation with Tinker. Similarly to the time spent at exhibit outcome of the earlier study, this outcome measures the ability of the agent to maintain the user engaged after having initiated the interaction.

6.3.4. Participants and Procedure. It was impractical to obtain consent from visitors who approached but then continued past the exhibit, especially given the 5,000 visitors a month who walked by. Furthermore, asking for consent prior to beginning the approach would have primed the subjects, invalidating the effects desired given that we were dealing with Tinker's first impressions. For these reasons, we collected data about *Number of Actions* during visitors' approaches and their *Session Duration* purely as anonymous behavioral outcomes.

After two weeks of pilot testing, we collected data from the 11th of December 2012 to the 15th of October 2013. Given the nature of the study, we did not collect demographics from study participants. However, the demographics of general museum visitors during the same time period indicate that most visitors are female (55%), aged 35–54 (57%), and well-educated (42% college degree and 41% graduate degree), and 77% visit with children.

Our experiment was administered in a completely automated fashion and some automated filtering rules were applied to the data collected. As soon as Tinker detected a visitor approaching, the visitor was randomized into one of the three conditions. Kendon's observations of human greetings were primarily focusing on 1-to-1 interactions and this impacted the design of our previous studies that had a single greeting agent and one user approaching it. In the museum setting, most visitors arrive in groups of two or more; therefore, we had to implement a series of filtering rules to pick social situations that fit our greeting model and obtain "clean" data by discarding from our dataset the cases shown in Table VI.

When more visitors were present Tinker started immediately to beckon them as if there was a single person that reached point T4. These cases were filtered out. When a single visitor that approached the exhibit and arrived at T4, the *"Encounter"* point, did not start the conversation session for a period longer than 4min Tinker

Table VI. The Filtering Rules Applied to Our Dataset. In Order to Obtain "Clean" Data We Discarded the Cases Indicated by the *Description* Column. The *Reason* Desribes Whether Theoretical or Study Design Assumptions were Violated

Filter #	Description	Reason
1	Approaches including more than one visitor	Violates theoretical assumptions
2	Single visitors walking backwards towards Tinker	Violates theoretical assumptions
3	Approaches started from intermediate points after the *"Distant Salutation"* (T1)	Violates theoretical assumptions
4	Visitors walking too fast at a speed above the threshold of 1m/s[a]	Stimuli not applicable
5	Visitors stopping while approaching[b]	Violates study assumptions
6	Returning visitors recognized by the hand reader	Violates study assumptions

[a]In earlier studies the approaching speed was 1.11m/s, but in the pilot testing at the museum we discovered that 1m/s was the maximum walking speed that allowed Tinker to exhibit the different nonverbal reactions considering its animation capabilities.
[b]A timeout of 8s started after hitting points [T1–T3] of the approach and Tinker started the beckon animation when it expired. The timeout was reset every time a point was reached.

would initiate the interaction. However, we included in our dataset cases when some passersby appeared from behind while a single visitor was involved in an ongoing "clean" approach.

In sum, we filtered out 15,441 cases out of a total of 30,727[3] visitors that were detected by Tinker during our data collection. Therefore, the total number of visitors across all conditions participating in our study and contributing to clean data is 15286. This number also includes visitors that performed *zero* actions and walked away while approaching Tinker (in a *clean* fashion).

6.3.5. Hypotheses. From the results of our preparatory studies, we predicted that:

—*H1 (Visitor Actions)*: Approaching visitors in the FRIENDLY TINKER group will perform a higher number of actions compared to the UNFRIENDLY TINKER group that, in turn, will complete a higher number of actions compared to the CONTROL group[4];
—*H2 (Session Duration)*: Given that a visitor has performed all the actions to complete the approach and start a conversation session with Tinker, visitors in the FRIENDLY TINKER sub-group will engage in longer sessions with the agent compared to the UNFRIENDLY TINKER sub-group that, in turn, will have longer sessions compared to the CONTROL sub-group.

6.3.6. Quantitative Analysis. Visitor Actions. A Kruskal Wallis test was conducted to compare this ordinal measure between the three groups. The test was non-significant ($\chi^2(2) = 1.58$, $p = 0.45$); therefore, *H1* was *not supported*. A further qualitative analysis of this outcome is discussed in the next section.

Session Duration. Duration had a positive skew; thus, we applied a $log_2 - transformation$ to the original dataset as recommended in Bland and Altman [1996]. We also detected outliers in our dataset corresponding to subjects who had a very short session duration. We opted to exclude those who had a duration less than to 10s, since this is the time required to exchange at least one dialogue turn with Tinker.

[3]N.b. this number includes also passers-by that came across the exhibit and quickly moved away.
[4]The CONTROL was expected to rate lowest because nothing was done to *catch the attention* (i.e. engage the visitor in interaction).

Table VII. Summary of the Visitor Actions Outcome. For each of the Three Experimental Conditions, the First Column Shows the Total Number of Cases. Then, from Left to Right, the Percentage of Visitors that Walked Away Prior to Performing any Action is Shown (in Parenthesis the Raw Number of Visitors), and in Succession the Number of Visitors that Performed the Action Described in the Heading Prior to Walking Away is Reported

Condition	Total # visitors	Visitor actions			
		Walked away	Approach completed	Hand in reader	Session started
Control	5085	89.99% (4576)	1.39% (71)	1.39% (71)	7.21% (367)
Unfriendly Tinker	5144	89.46% (4602)	1.34% (69)	1.78% (92)	7.40% (381)
Friendly Tinker	5057	89.28% (4515)	1.22% (62)	1.42% (72)	8.06% (408)

A D'Agostino skewness test [D'Agostino 1970] ran on the original dataset revealed a highly significant positive skewness of 2.6 ($SE = 14.07,\ p < 0.001$), whereas the $\log_2 -$ transformed data excluding outliers had a slight (non-significant) negative skewness of $-0.1\ (SE\ =\ -0.93,\ p\ =\ 0.35)$.

A one-way between-subjects ANOVA was conducted on the $\log_2 -$ transformed data to test for *Session Duration* differences among the three sub-groups that started a conversation with the agent ($N = 1138$, CONTROL = 365, UNFRIENDLY TINKER = 378, FRIENDLY TINKER = 395).

There were no significant main effects of study condition on *Session Duration*. In particular, the geometric mean[5] duration in the UNFRIENDLY TINKER group was higher ($GM\ =\ 80.50,\ 95\%\ CI\ [74.67, 86.77]$) compared to the FRIENDLY TINKER group ($GM = 77.92,\ 95\%\ CI\ [72.41, 83.88]$), that in turn, had higher geometric mean compared to the CONTROL GROUP ($GM\ =\ 73.77,\ 95\%\ CI\ [68.33, 79.62]$). However, this was not significant ($F(2, 1135) = 1.30$, $p = .27$); therefore, H2 was not supported.

6.3.7. Qualitative Observational Analysis. Since the hypotheses were not supported by our quantitative analysis, we exploited our data on Visitors Actions in a qualitative observational analysis described in this section, and an additional web survey in the next section aimed at understanding to what extent the manipulation of our field study was effective.

In Table VII, for each of the three experimental conditions, the first column shows the total number of cases in the given condition, and then the columns (from left to right) show the percentage of visitors (raw numbers are in parenthesis) that performed the corresponding action described in the heading prior to walking away, except for the first column that only shows how many visitors walked away without performing any action. For instance, if we examine the FRIENDLY TINKER condition, the row indicates that roughly the 90% of the visitors walked away, and then the 1.22% only completed the approach and then left, whereas another 1.42% of them went further by also placing the hand in the reader before walking away, and, finally, the 8.06% performed all of the actions by starting a conversation with Tinker.

These data would suggest that initial engagement worked better, encouraging visitors to start the session, when our model was active (7.40% and 8.06%) compared to the CONTROL condition (7.21%) where neither a greeting model nor impression management took place. When the model was present, FRIENDLY TINKER seemed

[5]A back transformation to the original scale of the *Session Duration* measure has been applied to the log-transformed means of the ANOVA test, this results in geometric and not arithmetic means. The confidence intervals (CI) shown are also back transformed.

Table VIII. The Retention Rates for Each Group. The Percentages Indicate Visitors that Tinker was Able to Retain in the Given Condition as They were Going Further with the Number of Actions Performed Prior to Walking Away

Retention rates				
		Visitor actions		
Condition	Total # visitors	Approach completed	Hand in reader	Session started
Control	5085	10.01%	8.62%	7.23%
Unfriendly Tinker	5144	10.54%	9.19%	7.41%
Friendly Tinker	5057	10.72%	9.50%	8.08%

to attract more visitors (8.06%) compared to the UNFRIENDLY version (7.40%) as originally predicted.

This can be seen more clearly in terms of retention rates as shown in Table VIII. For each condition, we computed the percentage of visitors that Tinker was able to retain after they performed each of the possible actions. We can see that Tinker, when managing first impressions, retained a higher number of visitors compared to the control version that was not initiating the interaction from distance. The FRIENDLY Tinker retained more visitors across the different actions compared to the UNFRIENDLY one.

The nonverbal greeting behavior exhibited by Tinker was critical for implementing our first impressions model and conduct our field study. It is possible that our manipulations failed and Tinker was not conveying to visitors the impressions we wanted, for this reason we went back in a controlled setting and (re)tested the difference between FRIENDLY and UNFRIENDLY TINKER versions in terms of friendliness judgments and first impressions that subjects had right after observing the two versions of the agent as presented in the following section.

6.3.8. First Impressions of Tinker in a Web Survey. Although we ran a manipulation check to selectively validate the nonverbal behavior exhibited by Tinker (i.e., smiling behavior and different amount of gaze at user) prior to conducting the field study, because of the importance of these stimuli, we questioned whether visitors were able to observe any difference between the two versions of Tinker that they had been exposed to.

Stimuli. We designed a within-subjects study where subjects accessed an on-line survey and observed in a fully randomized order two videos of Tinker, respectively the FRIENDLY and UNFRIENDLY versions, as if they were approaching it at the museum; thus, the agent was exhibiting the nonverbal reactions associated with each version in an automated fashion.

Measures. After watching each video, subjects were asked to rate their *Free Impressions* of the agent in the same format as described in Cafaro et al. [2012]. Thus, we asked them to write the first three adjectives that came to their minds (only one was mandatory). Then we asked them to report their impressions of *Tinker's Friendliness* by using a nine points Likert-scale ranging from *Extremely Unfriendly* to *Extremely Friendly*. This assessment of friendliness was also done in our behavior interpretation study and described in Cafaro et al. [2012].

Participants. We had 126 subjects participating in this on-line survey (78 males and 48 females). Participants were recruited via public announcements in our university campus and mailing lists. They represented six cultural identities[6] with 87% reporting

[6]We asked participants to select the nation that most represented their cultural identity from a list of all countries in the world as in earlier studies.

"Icelandic." They were aged 18–60 with 62% in the 21–30 range. Detailed demographics are provided in the electronic appendix.

Quantitative Analysis. Our hypothesis was that the FRIENDLY TINKER version would have received higher ratings of friendliness compared to the UNFRIENDLY TINKER one. We conducted a within-subjects one-way ANOVA on *Tinker Friendliness* ratings; however, there were not significant effects (($F(1, 125) = 1.41$, $p = 0.24$, $\eta_p^2 = 0.011$), thus we *rejected* our hypothesis.

Qualitative Analysis. We also performed a qualitative analysis of the *Free Impressions* adjectives following the methodology described in Cafaro et al. [2012]. We report here the total counts for the identified categories of adjectives with similar meanings in the following format (*<total count in the UNFRIENDLY TINKER condition>, <total count in the FRIENDLY TINKER condition>*). For the full dataset already grouped in categories, see the electronic appendix.

Tinker was judged as "weird, strange, awkward, and odd" (29 times vs. 26 times); "nice, polite, likable, and friendly" (21 times vs. 21 times); and "clever, assertive, and helpful" (12 times vs. 11 times) in both conditions.

UNFRIENDLY TINKER was more often judged as "mechanical, robotic, fake, and artificial" (31 times vs. 22 times) and also "lifeless, absent, dull, stiff, flat, and cold" (22 times vs. 14 times) compared to FRIENDLY TINKER.

On the other hand, the FRIENDLY TINKER version was more often seen as "funny, playful and childish" (7 times vs. 12 times), "silly, stupid, and useless" (17 times vs. 22 times), and considered more often "creepy, unnerving, and untrustworthy" (5 times vs. 20 times) than the other.

Regarding the graphical appearance of the agent, a few subjects also reported adjectives such as "evil eyes" and "crappy graphics" (for FRIENDLY TINKER) or "big eyes" (for UNFRIENDLY TINKER).

6.3.9. Discussion and Main Findings. The *Session Duration* assessment did not show significant improvements among the three versions of Tinker in our field study, though longer sessions were recorded when TINKER was managing impressions compared to the CONTROL version. We believe that dialogue content affect this outcome more than Tinker's greeting behavior displayed while initiating the interaction. It is possible that by adding a brief spoken greeting speech—manipulated according to the two levels of friendliness—could have had an impact on this assessment, though our focus was on nonverbal cues prior to beginning any verbal exchange. It should be emphasized that the earlier field study got significant improvements on a similar outcome (i.e., Time Spent at Exhibit) when relational verbal and nonverbal behavior was exhibited. During our field study Tinker always had the relational behaviors turned on, which might explain why no consistent differences, other than the small trends, were found.

The qualitative analysis of *Visitor Actions* revealed an important finding, turning on our first impression model was attracting more visitors to the exhibit, more specifically when Tinker managed friendly impressions. According to Richards [2012], virtual guides must include an agent wait state. These are the believable behaviors that the agent should perform when they are not interacting with a user. This is quite common in exhibition guides. Often this wait state is used to attract the attention of passersby at an exhibit. Thus, our results complied with these suggestions.

The *web survey* demonstrated that subjects did not observe any difference in terms of friendliness between the two videos of Tinker that they observed. Although the UNFRIENDLY TINKER version was judged as more "mechanical, robotic, fake, and artificial" than the FRIENDLY TINKER, and we think this might be due to the lack of nonverbal cues expressed by the former (i.e., no smiling and a fewer gazes at the visitor), the two versions of Tinker were judged equally as “nice, polite, likable, and friendly.” By analyzing subjects’ responses in the web survey, we noticed that the smiling behavior of the FRIENDLY TINKER version led subjects to report adjectives such as “evil eyes” or “Tinker is over friendly, the agent seems over excited”.

It is interesting to note that FRIENDLY TINKER was more often judged as “silly, stupid, and useless” in the web survey, this might be related to the trends showing visitors having longer session duration with the UNFRIENDLY TINKER. Furthermore, FRIENDLY TINKER was considered “untrustworthy.” This might have important implications in terms of the agent’s pedagogical and relational role, since visitors might have trusted the UNFRIENDLY TINKER version more when it came to asking information or directions about the museum exhibits and bond with the agent.

From this follow-up web study and our main field study, we can conclude that extending Tinker with our model of first impression seemed to encourage more visitors to engage in interaction, but in order to obtain a greater substantial effect more factors need to be considered. First, greeting behavior—smiling in particular—assumes a different meaning in different social contexts and settings. The effects on participants were different when moving from the controlled to the public space setting, and exhibiting this behavior, in such a public setting, becomes trickier. Secondly, the particular animation for this behavior on the specific agent model at the museum (i.e., Tinker) led some subjects to judge the agent as untrustworthy and silly, whereas our intent was to create an interface for visitors that would make impressions of being friendly and welcoming. Thirdly, the large number of visitors that walked away prior to interacting with Tinker could be independent from the behaviors that were displayed by the agent. In such public space, it is not unusual that visitors shop around for interesting exhibits and often limit their engagement (i.e., approaching and time spent with an exhibit) to a cursory glance. A previous observational study about visitors’ behavior in a science museum revealed that individuals not in a family group spent, on average, less than 1min per exhibit both during weekdays and weekends [Sandifer 1997].

The noisy environment of the museum is another limitation of our study. Surrounding people and activities can distract subjects and provide competing stimuli, e.g., causing them to miss important behavior.

6.4. Impact of Results on First Impressions Model

From this final field and web studies we learned that first impressions of friendliness managed along a greeting model for initiating the interaction help attaining more initial engagement, in our case, towards an anthropomorphic virtual agent.

Starting with our initial proposed model we found that migrating from the laboratory to a real life setting there are possible variations of the selected social theories that need to be considered. We were able to retain the original greeting model with filtering and manage the impression of attitude through smiling and gazing immediacy behavior. However, there are several aspects to be considered for improving our model and possibly achieve stronger results both on initial engagement and retention.

We believe that *Expectancy Violation Theory*, introduced by Burgoon and Hale [1988], provides a valuable theoretical background for discussing those aspects. EVT proposes that observation and interaction with others leads to expectancies; it can apply to both non-relational interaction and close relationships. According to this theory, five core factors affect expectancies: *environmental variables*, *interaction variables*, *context*, *communicator characteristics,* and *relational characteristics* [Griffin 2012, chapter 7]. This theory suggests ways our studies could have been better controlled, as well directions for future research. We included personality in our model (a communicator characteristic); however, there are interesting aspects related to appearance that open to interesting extensions. Agent-side these are (1) the physical embodiment (e.g.

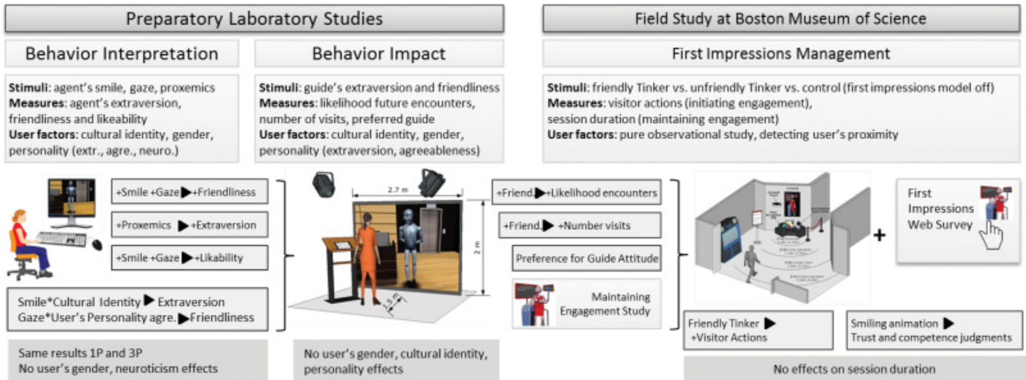

Fig. 6. A summary of our results. The field study on impression management was informed by results of two preparatory studies shown on the left, and a former field study on maintaining engagement conducted at the museum. The final results obtained in the field and from the web survey are depicted on the right. Main effect directions are marked with "+" (e.g., more gaze accounted for a higher level of agent friendliness). Interaction effects are marked with "*."

virtual vs. robotic), (2) the graphical appearance when virtual (e.g. robotic, human), (3) the level of realism, and (4) the appeal (e.g. clothing style). On the user side, we have aspects to consider, such as gender, cultural identity, sex, and age. The virtual embodiment is more flexible when it comes to manipulating visual characteristics (e.g., appeal, level of realism). A different appearance, for example, may impact the behavior realization (i.e. animation), thus resulting in different animation styles and interpretations of the same behavior, for example, of the smiling behavior between our controlled and field studies. Several researchers have studied the effects of these aspects on user's perception of virtual characters [McDonnell et al. 2012; Zibrek and McDonnell 2014]. Museum visitors in the field study might have arrived at the exhibit area with different expectations. They may think of the agent as an artificial object of curiosity rather than a living being and therefore not expect socially appropriate greeting behavior, and the manipulation of these expectations can be explored in future studies. Regarding context, the transition from the monitor setting of the first laboratory study to the life-sized display of Tinker at the museum could have influenced the overall user experience. In our second laboratory study on behavior impact, we used a life-sized version of our agent in order to make this transition as smooth as possible toward the field study and for displaying our stimuli better (e.g., smiling facial expressions).

7. CONCLUSIONS

We presented a first impressions model for initiating user–agent interactions in the context of a first greeting encounter. The model extends the capabilities of an anthropomorphic interface agent with impression management of personality (extraversion) and interpersonal attitude (affiliation) through subtle manipulations of nonverbal immediacy cues (i.e., smiling, gazing, and proxemics behavior) exhibited during the initial moments of interaction.

We presented two laboratory studies that implemented our model and informed the design of a field study at the Boston MoS where we evaluated the effects of managing first impressions on the initial engagement (i.e., attraction to begin interaction) of Tinker, a relational guide agent.

An overview of our results is depicted in Figure 6.

In the *behavior interpretation* study, we showed that users quickly form *distinct* impressions of extraversion when greeting agents exhibit proxemics behavior (i.e., stepped toward the user) and judgments of friendliness are made when agents smile and gaze more at the users. Therefore, subtle manipulations of immediacy cues can be used to manage, simultaneously, personality and attitude of a greeting agent. We demonstrated that camera perspective had no influence on user interpretations of an agent's nonverbal behavior given the interaction was limited to just observing the approaches. Likability of the agent (i.e., willingness to spend more time with it) mirrored results on friendliness. User's cultural identity and personality interacted with some of the behavior exhibited. In particular, we discovered that only those scoring low in agreeableness interpreted agents gazing more at them as friendlier compared to agents gazing less. The subjects' gender and neuroticism trait did not show any impact on the behavior interpretation.

In the *behavior impact* study, we found that the nonverbal greeting behavior exhibited by life-sized agents had impact on subjects' relational decisions in terms of how likely and how often they would like to spend time with the agents on future virtual guided tours of a 3D virtual museum. We found that agents' interpersonal attitude had greater impact compared to personality on subjects' decisions, in particular, the agents exhibiting high friendliness behavior were more likely to be encountered again and more frequently than the ones exhibiting low friendliness, no matter the personality expressed. Those guides managing their friendliness (attitude) got the most preferences, even though extroverted guides (personality) where shown. We did not observe effects of subject's own personality, gender, and cultural identity.

In the *field* study at the MoS, we detected visitors while approaching Tinker, a virtual relational museum guide exhibit. A former field study focused on maintaining engagement with Tinker through relational behavior. Building on this and our preparatory studies, we implemented our first impressions model to evaluate whether managing friendliness led to a better initial engagement encouraging visitors to initiate an interaction with the agent. Observational results showed that more visitors initiated an interaction with Tinker when our model was working (i.e. turned on), and a friendly version of the agent attracted more visitors compared to the unfriendly one.

In conclusion, we demonstrated that *first impressions* of *personality* and *interpersonal attitude* based on quick judgments of observed nonverbal behavior are important for conversational agents, in general, but particularly for those in public spaces. We showed that the particular *nonverbal immediacy cues* of *smile, gaze, and proxemics* during a first user–agent encounter have effect on user's interpretations in terms of agent's extraversion and affiliation, and have impact on users' relational decisions in terms of likelihood and frequency of further encounters.

Finally, we discovered that the communicative greeting models chosen and the concept of interpersonal distance between user and agent are a valuable means for defining the user–agent interaction. Anthropomorphic user interfaces can benefit from these findings. In particular, in public deployments such as welcoming museum visitors for guided tours, extending these agents with our first impressions model and incorporating a greeting theory can improve initial engagement to the exhibit.

8. LIMITATIONS AND FUTURE WORK

Some limitations of our approach should be considered. Kendon's greeting model describes 1-to-1 human–human interactions; thus we were only able to model single user–agent greeting encounters. However, in the museum groups of people often approach the exhibits. Further exploration is needed with different configurations. Huang and colleagues [Huang et al. 2010] were able to express different attitudes in multiparty agent–users interaction while exhibiting idle behavior. Dim and Kuflik [2014] tracked social behavior (e.g., proximity, orientation towards exhibits) of pairs in a museum to deduce interest in the exhibits. Moreover, the theoretical foundations of our first impression model are based on social studies conducted with a sample population belonging to "Western cultures;" therefore, the results presented in this work are more likely to represent a Western contemporary context.

It would be good to better understand the role of the user's own personality when interpreting an agent's nonverbal behavior and expressing a preference for a particular agent type. We may also want to look further into detection and influences of other factors (e.g. cultural identity, age) and their interaction with personality.

It would be interesting to perform a longitudinal study that would allow us to quantitatively assess technology acceptance, and examine the stability and long-term impact of the impressions formed in the brief user-agent encounters, i.e. how long it takes to overcome those impressions. In such design, a behavioral measure aimed at assessing whether users actually return to interact with an RA can be assessed.

The studies presented in this paper implement a first impression model that lends itself to new intriguing lines of research. The agent's verbal behavior and more nonverbal cues can be manipulated, and impressions of different personality traits (e.g. neuroticism) and attitude (e.g. status dimension) can be considered. In addition to personality and interpersonal attitude, it could be relevant for an RA to manage impressions of *competence* or *trustworthiness*. Some recent work on modeling interpersonal trust in social robots has been done [Lee et al. 2013], but further research in the context of first greeting encounters is still needed. Finally, it can be interesting to deploy our first impressions model and focus on manipulating other (visual) agent's characteristics, such as embodiment (virtual vs. robotic), appeal (e.g. cloth style), and level of realism.

ELECTRONIC APPENDIX

The electronic appendix for this article can be accessed in the ACM Digital Library.

ACKNOWLEDGMENTS

We are grateful to the Boston's Museum of Science and the staff at Cahner's Computer Place for their assistance, and to Juan Fernandez for maintaining Tinker and providing help with the system setup and data collection.

REFERENCES

Samer Al Moubayed, Jonas Beskow, Gabriel Skantze, and Björn Granström. 2012. Furhat: A back-projected human-like robot head for multiparty human-machine interaction. In *Cognitive Behavioural Systems*, Anna Esposito, Antonietta M. Esposito, Alessandro Vinciarelli, Rüdiger Hoffmann, and Vincent C. Müller (Eds.), Lecture Notes in Computer Science, Vol. 7403. Springer Berlin Heidelberg, 114–130. DOI:http://dx.doi.org/10.1007/978-3-642-34584-5_9

I. Altman and D. Taylor. 1973. *Social Penetration: The Development of Interpersonal Relationships*. Holt, Rinhart & Winston, New York, NY.

Nalini Ambady and Robert Rosenthal. 1993. Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *J. Personal. Soc. Psychol.* 64, 3 (1993), 431–441. DOI:http://dx.doi.org/10.1037/0022-3514.64.3.431

Nalini Ambady and John J. Skowronski. 2008. *First Impressions*. Guilford Press, New York, NY. 368 pages.

Elisabeth André, Martin Klesen, Patrick Gebhard, Steve Allen, and Thomas Rist. 2000. Chapter integrating models of personality and emotions into lifelike characters, In *Affective Interactions*. Springer-Verlag New York, Inc., New York, NY, 150–165. http://dl.acm.org/citation.cfm?id=362587.374273

Michael Argyle. 1988. *Bodily Communication* (2nd ed.). Methuen, New York. 363 pages.

Jens B. Asendorpf and Susanne Wilpers. 1998. Personality effects on social relationships. *J. Personal. Soc. Psychol.* 74, 6 (1998), 1531–1544. DOI:http://dx.doi.org/10.1037/0022-3514.74.6.1531

Daniel Ballin, Marco Gillies, and Barry Crabtree. 2004. A framework for interpersonal attitude and nonverbal communication in improvisational visual media production. In *Proceedings of the First European Conference on Visual Media Production (CVMP) (CVMP)*. IEEE Press, 203–210.

Moshe Bar, Maital Neta, and Heather Linz. 2006. Very first impressions. *Emotion* 6, 2 (2006), 269.

Tim Barnett, Allison W. Pearson, Rodney Pearson, and Franz W. Kellermanns. 2015. Five-factor model personality traits as predictors of perceived and actual usage of technology. *European Journal of Information Systems* 24, 4 (2015), 374–390. DOI:http://dx.doi.org/10.1057/ejis.2014.10

Kirsten Bergmann, Friederike Eyssel, and Stefan Kopp. 2012. A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In *Intelligent Virtual Agents*, Yukiko Nakano, Michael Neff, Ana Paiva, and Marilyn Walker (Eds.), Lecture Notes in Computer Science, Vol. 7502. Springer-Verlag, Berlin, Heidelberg, 126–138. DOI:http://dx.doi.org/10.1007/978-3-642-33197-8_13

Timothy Bickmore and Justine Cassell. 2001. Relational agents: A model and implementation of building user trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'01)*. ACM, New York, NY, 396–403. DOI:http://dx.doi.org/10.1145/365024.365304

Timothy Bickmore and Rosalind Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Trans. Comput. Hum. Interact.* 12, 2 (2005), 293–327. DOI:http://dx.doi.org/10.1145/1067860.1067867

Timothy W. Bickmore, Lisa Caruso, Kerri Clough-Gorr, and Tim Heeren. 2005. 'It's just like you talk to a friend' relational agents for older adults. *Interact. Comput.* 17, 6 (Dec. 2005), 711–735. DOI:http://dx.doi.org/10.1016/j.intcom.2005.09.002

Timothy W. Bickmore, Suzanne E. Mitchell, Brian W. Jack, Michael K. Paasche-Orlow, Laura M. Pfeifer, and Julie O'Donnell. 2010. Response to a relational agent by hospital patients with depressive symptoms. *Interact. Comput.* 22, 4 (July 2010), 289–298. DOI:http://dx.doi.org/10.1016/j.intcom.2009.12.001

Timothy W. Bickmore, Laura Pfeifer, Daniel Schulman, Sepalika Perera, Chaamari Senanayake, and Ishraque Nazmi. 2008. Public displays of affect: Deploying relational agents in public spaces. In *Proceedings of the CHI'08 Extended Abstracts on Human Factors in Computing Systems (CHI EA'08)*. ACM, New York, NY, 3297–3302. DOI:http://dx.doi.org/10.1145/1358628.1358847

Timothy W. Bickmore, Laura M. Pfeifer, and Brian W. Jack. 2009. Taking the time to care: Empowering low health literacy hospital patients with virtual nurse agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*. ACM, New York, NY, 1265–1274. DOI:http://dx.doi.org/10.1145/1518701.1518891

Timothy W. Bickmore, Laura M. Pfeifer Vardoulakis, and Daniel Schulman. 2013. Tinker: A relational agent museum guide. *Autonom. Agents Multi-Agent Syst.* 27, 2 (2013), 254–276. DOI:http://dx.doi.org/10.1007/s10458-012-9216-7

Timothy W. Bickmore, Kathryn Puskar, Elizabeth A. Schlenk, Laura M. Pfeifer, and Susan M. Sereika. 2010. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interact. Comput.* 22, 4 (July 2010), 276–288. DOI:http://dx.doi.org/10.1016/j.intcom.2010.02.001

Martin J. Bland and Douglas Altman G. 1996. Statistics notes: Transforming data. *BMJ* 312, 7033 (1996), 770. DOI:http://dx.doi.org/10.1136/bmj.312.7033.770

Kenneth S. Bordens and Bruce B. Abbott. 2002. *Research Design and Methods: A Process Approach* (8th ed. ed.). McGraw-Hill, New York, NY. 624 pages.

Peter Borkenau, Nadine Mauer, Rainer Riemann, Frank M Spinath, and Alois Angleitner. 2004. Thin slices of behavior as cues of personality and intelligence. *J. Personal. Soc. Psychol.* 86, 4 (2004), 599. DOI:http://dx.doi.org/10.1037/0022-3514.86.4.599

Judee K. Burgoon, David B. Buller, Jerold L. Hale, and Mark A. de Turck. 1984. Relational messages associated with nonverbal behaviors. *Hum. Commun. Res.* 10, 3 (1984), 351–378. DOI:http://dx.doi.org/10.1111/j.1468-2958.1984.tb00023.x

Judee K. Burgoon and Jerold L. Hale. 1984. The fundamental topoi of relational communication. *Commun. Monogr.* 51, 3 (1984), 193–214.

Judee K. Burgoon and Jerold L. Hale. 1988. Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors. *Commun. Monogr.* 55, 1 (1988), 58–79. DOI:http://dx.doi.org/10.1080/03637758809376158

Angelo Cafaro, Hannes H. Vilhjálmsson, Timothy W. Bickmore, Dirk Heylen, and Daniel Schulman. 2013. First impressions in user-agent encounters: The impact of an agent's nonverbal behavior on users' relational decisions. In *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS'13)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1201–1202. http://dl.acm.org/citation.cfm?id=2484920.2485142

Angelo Cafaro, HannesHögni Vilhjálmsson, Timothy Bickmore, Dirk Heylen, KamillaRún Jóhannsdóttir, and GunnarSteinn Valgarðsson. 2012. First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters. In *Intelligent Virtual Agents*, Yukiko Nakano, Michael Neff, Ana Paiva, and Marilyn Walker (Eds.), Lecture Notes in Computer Science, Vol. 7502. Springer Berlin Heidelberg, Richland, SC, 67–80. DOI:http://dx.doi.org/10.1007/978-3-642-33197-8_7

Anne Campbell and J. Philippe Rushton. 1978. Bodily communication and personality. *Br. J. Soc. Clin. Psychol.* 17, 1 (1978), 31–36. DOI:http://dx.doi.org/10.1111/j.2044-8260.1978.tb00893.x

J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjálmsson, and H. Yan. 1999. Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'99)*. ACM, New York, NY, 520–527. DOI:http://dx.doi.org/10.1145/302979.303150

J. Cassell, T. Bickmore, H. Vilhjálmsson, and H. Yan. 2000. More than just a pretty face: Affordances of embodiment. In *Proceedings of the 5th International Conference on Intelligent user Interfaces (IUI'00)*. ACM, New York, NY, 52–59. DOI:http://dx.doi.org/10.1145/325737.325781

J. Cassell, T. Stocky, T. Bickmore, Y. Gao, Y. Nakano, K. Ryokai, C. Tversky, D.and Vaucelle, and H. VilhjÁlmsson. 2002. MACK: Media lab autonomous conversational kiosk. In *Proceedings of Imagina'02*. 12–15.

Ginevra Castellano, Ana Paiva, Arvid Kappas, Ruth Aylett, Helen Hastie, Wolmet Barendregt, Fernando Nabais, and Susan Bull. 2013. Towards empathic virtual and robotic tutors. In *Artificial Intelligence in Education*, H. Chad Lane, Kalina Yacef, Jack Mostow, and Philip Pavlik (Eds.), Lecture Notes in Computer Science, Vol. 7926. Springer Berlin Heidelberg, 733–736. DOI:http://dx.doi.org/10.1007/978-3-642-39112-5_100

Raymond B. Cattell. 1965. *The Scientific Analysis of Personality*. Penguin Books, Baltimore, MD. 399 pages.

Aleksandra Cerekovic, Oya Aran, and Daniel Gatica-Perez. 2014. How do you like your virtual agent?: Human-agent interaction experience through nonverbal features and personality traits. In *Human Behavior Understanding*, HyunSoo Park, AlbertAli Salah, YongJae Lee, Louis-Philippe Morency, Yaser Sheikh, and Rita Cucchiara (Eds.), Lecture Notes in Computer Science, Vol. 8749. Springer International Publishing, 1–15. DOI:http://dx.doi.org/10.1007/978-3-319-11839-0_1

Mathieu Chollet, Magalie Ochs, and Catherine Pelachaud. 2014. From non-verbal signals sequence mining to bayesian networks for interpersonal attitudes expression. In *Intelligent Virtual Agents*, Timothy Bickmore, Stacy Marsella, and Candace Sidner (Eds.), Lecture Notes in Computer Science, Vol. 8637. Springer International Publishing, 120–133. DOI:http://dx.doi.org/10.1007/978-3-319-09767-1_15

Tim Cole and James J. Bradac. 1996. A lay theory of relational satisfaction with best friends. *J. Soc. Personal Relationsh.* 13, 1 (1996), 57–83. DOI:http://dx.doi.org/10.1177/0265407596131004

Ralph B. D'Agostino. 1970. Transformation to normality of the null distribution of g1. *Biometrika* 57, 3 (1970), 679–681. DOI:http://dx.doi.org/10.1093/biomet/57.3.679

Fred D. Davis. 1993. User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *Int. J. Man-Mach. Stud.* 38, 3 (March 1993), 475–487. DOI:http://dx.doi.org/10.1006/imms.1993.1022

Zhigang Deng, Jeremy Bailenson, J. P. Lewis, and Ulrich Neumann. 2006. Perceiving visual emotions with speech. In *Intelligent Virtual Agents*, Jonathan Gratch, Michael Young, Ruth Aylett, Daniel Ballin, and Patrick Olivier (Eds.), Lecture Notes in Computer Science, Vol. 4133. Springer Berlin Heidelberg, 107–120. DOI:http://dx.doi.org/10.1007/11821830_9

Bella M. DePaulo. 1992. Nonverbal behavior and self-presentation. *Psychol. Bull.* 111, 2 (1992), 203–243. DOI:http://dx.doi.org/10.1037//0033-2909.111.2.203

Amol Deshmukh, Ginevra Castellano, Mei Yii Lim, Ruth Aylett, and Peter W. McOwan. 2010. Ubiquitous social perception abilities for interaction initiation in human-robot interaction. In *Proceedings of the 3rd International Workshop on Affective Interaction in Natural Environments (AFFINE'10)*. ACM, New York, NY, 55–60. DOI:http://dx.doi.org/10.1145/1877826.1877841

Sarv Devaraj, Robert F. Easley, and J. Michael Crant. 2008. Research note'how does personality matter? Relating the five-factor model to technology acceptance and use. *Inf. Syst. Res.* 19, 1 (2008), 93–105. DOI:http://dx.doi.org/10.1287/isre.1070.0153

Eyal Dim and Tsvi Kuflik. 2014. Automatic detection of social behavior of museum visitor pairs. *ACM Trans. Interact. Intell. Syst.* 4, 4, Article 17 (Nov. 2014), 30 pages. DOI:http://dx.doi.org/10.1145/2662869

Tiago Doce, João Dias, Rui Prada, and Ana Paiva. 2010. Creating individual agents through personality traits. In *Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.), Lecture Notes in Computer Science, Vol. 6356. Springer Berlin Heidelberg, 257–264. DOI:http://dx.doi.org/10.1007/978-3-642-15892-6_27

P. Ekman and W. Friesen. 1978. *The Facial Action Coding System: A Technique For The Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, CA.

H. J. Eysenck. 1953. The scientific study of personality. *Br. J. Stat. Psychol.* 6, 1 (1953), 44–52. DOI:http://dx.doi.org/10.1111/j.2044-8317.1953.tb00132.x

Atsushi Fukayama, Takehiko Ohno, Naoki Mukawa, Minako Sawaki, and Norihiro Hagita. 2002. Messages embedded in gaze of interface agents—Impression management with agent's gaze. In *Proceedings of* *the SIGCHI Conference on Human Factors in Computing Systems (CHI'02)*. ACM, New York, NY, USA, 41–48. DOI:http://dx.doi.org/10.1145/503376.503385

Andrew Gelman and David Park. 2009. Splitting a predictor at the upper quarter or third and the lower quarter or third. *Am. Stat.* 63, 1 (2009), 1–8.

C. Gelso and J. Hayes. 1998. *The Psychotherapy Relationship: Theory, Research and Practice*. John Wiley and Sons, New York, NY.

Madhumita Ghosh and Hideaki Kuzuoka. 2014. An ethnomethodological study of a museum guide robot's attempt at engagement and disengagement. *J. Robot.* 2014 (2014). DOI:http://dx.doi.org/10.1155/2014/876439

Debra S. Gill, Andrew Christensen, and Fransk D. Fincham. 1999. Predicting marital satisfaction from behavior: Do all roads really lead to rome? *Personal Relationsh.* 6, 3 (1999), 369–387. DOI:http://dx.doi.org/10.1111/j.1475-6811.1999.tb00198.x

Rachel Gockley, Allison Bruce, Jodi Forlizzi, Marek Michalowski, Anne Mundell, Stephanie Rosenthal, Brennan Sellner, Reid Simmons, Kevin Snipes, Alan Schultz, and Jue Wang. 2005. Designing robots for long-term social interaction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005*. 1338–1343. DOI:http://dx.doi.org/10.1109/IROS.2005.1545303

Erving Goffman. 1959. *The Presentation of Self in Everyday Life* (1st ed.). Doubleday, Garden City, New York, NY. 259 pages.

Paulo F. Gomes, Alberto Sardinha, Elena Márquez Segura, Henriette Cramer, and Ana Paiva. 2014. Migration between two embodiments of an artificial pet. *In. J. Humanoid Robot.* 11, 01 (2014), 1450001. DOI:http://dx.doi.org/10.1142/S0219843614500017

Em Griffin. 2012. *A First Look at Communication Theory* (8th ed.). McGraw-Hill Companies, Inc. 84–92 pages.

Edward T. Hall. 1966. *The Hidden Dimension*. Doubleday.

L. Havens. 1986. *Making Contact: Uses of Language in Psychotherapy*. Harvard University Press, Cambridge, MA.

Brandon Heenan, Saul Greenberg, Setareh Aghelmanesh, and Ehud Sharlin. 2013. *Designing Social Greetings and Proxemics in Human Robot Interaction*. Technical Report. Department of Computer Science, University of Calgary, Calgary, AB, Canada.

Hung-Hsuan Huang, Takuya Furukawa, Hiroki Ohashi, Toyoaki Nishida, Aleksandra Cerekovic, Igor S. Pandzic, and Yukiko Nakano. 2010. Communicating with multiple users for embodied conversational agents in quiz game context. *Int. J. Intell. Inf. Database Syst.* 4, 6 (Nov. 2010), 579–598. DOI:http://dx.doi.org/10.1504/IJIIDS.2010.036896

Katherine Isbister and Clifford Nass. 2000. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *Int. J. Hum.-Comput. Stud.* 53, 2 (2000), 251–267. DOI:http://dx.doi.org/10.1006/ijhc.2000.0368

Lara K. Kammrath, Daniel R. Ames, and Abigail A. Scholer. 2007. Keeping up impressions: Inferential rules for impression change across the big five. *J. Exp. Soc. Psychol.* 43, 3 (2007), 450–457. DOI:http://dx.doi.org/10.1016/j.jesp.2006.04.006

Sin-Hwa Kang, Jonathan Gratch, Candy Sidner, Ron Artstein, Lixing Huang, and Louis-Philippe Morency. 2012. Towards building a virtual counselor: Modeling nonverbal behavior during intimate self-disclosure. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 63–70.

Sin-Hwa Kang, Jonathan Gratch, Ning Wang, and James H. Watt. 2008a. Agreeable people like agreeable virtual humans. In *Proceedings of the 8th International Conference on Intelligent Virtual Agents*. Lecture Notes in Computer Science, Springer-Verlag, Berlin, 253–261. DOI:http://dx.doi.org/10.1007/978-3-540-85483-8_26

Sin-Hwa Kang, Jonathan Gratch, Ning Wang, and James H. Watt. 2008b. Does the contingency of agents' nonverbal feedback affect users' social anxiety? In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 120–127.

Jack Katz. 2001. *How Emotions Work*. University of Chicago Press.

Adam Kendon. 1990. *Conducting Interaction: Patterns of Behavior in Focused Encounters (Studies in Interactional Sociolinguistics)*. Cambridge University Press, New York, NY.

Yunkyung Kim and Bilge Mutlu. 2014. How social distance shapes human–robot interaction. *Int. J. Hum.-Comput. Stud.* 72, 12 (2014), 783–795. DOI:http://dx.doi.org/10.1016/j.ijhcs.2014.05.005

Kheng Lee Koay, DagSverre Syrdal, Mohammadreza Ashgari-Oskoei, MichaelL. Walters, and Kerstin Dautenhahn. 2014. Social roles and baseline proxemic preferences for a domestic service robot. *Int. J. Soc. Robot.* 6, 4 (2014), 469–488. DOI:http://dx.doi.org/10.1007/s12369-014-0232-4

Stefan Kopp, Lars Gesellensetter, Nicole C. Krämer, and Ipke Wachsmuth. 2005. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, Themis Panayiotopoulos, Jonathan Gratch, Ruth Aylett, Daniel Ballin, Patrick Olivier, and Thomas Rist (Eds.), Lecture Notes in Computer Science, Vol. 3661. Springer Berlin Heidelberg, 329–343. DOI:http://dx.doi.org/10.1007/11550617_28

Brigitte Krenn, Birgit Endrass, Felix Kistler, and Elisabeth André. 2014. Effects of language variety on personality perception in embodied conversational agents. In *Human-Computer Interaction. Advanced Interaction Modalities and Techniques*. Masaaki Kurosu (Ed.), Lecture Notes in Computer Science, Vol. 8511. Springer International Publishing, 429–439. DOI:http://dx.doi.org/10.1007/978-3-319-07230-2_41

H. Chad Lane, Dan Noren, Daniel Auerbach, Mike Birch, and William Swartout. 2011. Intelligent tutoring goes to the museum in the big city: A pedagogical agent for informal science education. In *Artificial Intelligence in Education*, Gautam Biswas, Susan Bull, Judy Kay, and Antonija Mitrovic (Eds.), Lecture Notes in Computer Science, Vol. 6738. Springer Berlin Heidelberg, 155–162. DOI:http://dx.doi.org/10.1007/978-3-642-21869-9_22

John Laver. 1981. Linguistic routines and politeness in greeting and parting. In *Conversational Routine*, Florian Coulmas (Ed), Mouton, New York, NY, 289–304.

Jina Lee and Stacy Marsella. 2011. Modeling side participants and bystanders: The importance of being a laugh track. In *Intelligent Virtual Agents*, Hannes H.i Vilhjálmsson, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson (Eds.), Lecture Notes in Computer Science, Vol. 6895. Springer Berlin Heidelberg, 240–247. DOI:http://dx.doi.org/ 10.1007/978-3-642-23974-8_26

Jin Joo Lee, Brad Knox, Jolie Baumann, Cynthia Breazeal, and David DeSteno. 2013. Computationally modeling interpersonal trust. *Front. Psychol.* 4, 893 (2013). DOI:http://dx.doi.org/10.3389/fpsyg.2013.00893

Maurice J. Levesque and David A. Kenny. 1993. Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *J. Personal. Soc. Psychol.* 65, 6 (1993), 1178–1187. DOI:http://dx.doi.org/10.1198/tast.2009.0001

Mei Yii Lim, Michael Kriegel, Ruth Aylett, and Amol Deshmukh. 2010. Sarah: An ECA in the wild. In *Proceedings of the International Workshop on Interacting with ECAs as Virtual Characters*. 53.

Kris Liu, Jackson Tolins, Jean E. Fox Tree, Marilyn Walker, and Michael Neff. 2013. Judging IVA personality using an open-ended question. In *Intelligent Virtual Agents*, Ruth Aylett, Brigitte Krenn, Catherine Pelachaud, and Hiroshi Shimodaira (Eds.), Lecture Notes in Computer Science, Vol. 8108. Springer Berlin Heidelberg, 396–405. DOI:http://dx.doi.org/10.1007/978-3-642-40415-3_23

Gale M. Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. 2014. It's only a computer: Virtual humans increase willingness to disclose. *Comput. Hum. Behav.* 37 (Aug. 2014), 94–100. DOI:http://dx.doi.org/10.1016/j.chb.2014.04.043

Y. Ludewig, N. Doring, and N. Exner. 2012. Design and evaluation of the personality trait extraversion of a shopping robot. In *Proceedings of the RO-MAN, 2012 IEEE*. 372–379. DOI:http://dx.doi.org/10.1109/ROMAN.2012.6343781

Mark Maat and Dirk Heylen. 2009. Turn management or impression management? In *Proceedings of the 9th International Conference on Intelligent Virtual Agents (IVA'09)*. Springer-Verlag, Berlin, Heidelberg, 467–473. DOI:http://dx.doi.org/10.1007/978-3-642-04380-2_51

Mark Ter Maat, Khiet P. Truong, and Dirk Heylen. 2010. How turn-taking strategies influence users' impressions of an agent. In *Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.), Lecture Notes in Computer Science, Vol. 6356. Springer Berlin Heidelberg, 441–453. DOI:http://dx.doi.org/10.1007/978-3-642-15892-6_48

Robert MacCallum, Shaobo Zhang, Kristopher Preacher, and Derek Rucker. 2002. On the practice of dichotomization of quantitative variables. *Psychol. Methods* 7, 1 (2002), 19–40. DOI:http://dx.doi.org/10.1037/1082-989X.7.1.19

Robert R. McCrae and Jr. Costa, Paul T. 1997. Personality trait structure as a human universal. *Am. Psychol.* 52, 5 (1997), 509–516. DOI:http://dx.doi.org/10.1037/0003-066X.52.5.509

Rachel McDonnell, Martin Breidt, and Heinrich H. Bülthoff. 2012. Render me real?: Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graph.* 31, 4, Article 91 (July 2012), 11 pages. DOI:http://dx.doi.org/10.1145/2185520.2185587

Marc Mehu, Anthony C. Little, and Robin I.M. Dunbar. 2008. Sex differences in the effect of smiling on social judgments: An evolutionary approach. *J. Soc., Evol. Cult. Psychol.* 2, 3 (2008), 103–121. DOI:http://dx.doi.org/10.1037/h0099351

Rowland Miller, Daniel Perlman, and Sharon Brehm. 2007. *Intimate Relationships* (4th ed.). McGraw-Hill, Boston, MA.

Clifford Nass, Youngme Moon, B. J. Fogg, Byron Reeves, and Chris Dryer. 1995. Can computer personalities be human personalities? In *Proceedings of the Conference Companion on Human Factors in Computing Systems (CHI'95)*. ACM, New York, NY, USA, 228–229. DOI:http://dx.doi.org/10.1145/223355.223538

L. P. Naumann, S. Vazire, P. J. Rentfrow, and S. D. Gosling. 2009. Personality judgments based on physical appearance. *Personal. Soc. Psychol. Bull.* 35, 12 (2009), 1661–1671. DOI:http://dx.doi.org/10.1177/0146167209346309

Michael Neff, Nicholas Toothman, Robeson Bowmani, Jean E. Fox Tree, and Marilyn A. Walker. 2011. Don't scratch! self-adaptors reflect emotional stability. In *Intelligent Virtual Agents*, Hannes H. VilhjÁlmsson, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson (Eds.), Lecture Notes in Computer Science, Vol. 6895. Springer Berlin Heidelberg, 398–411. DOI:http://dx.doi.org/ 10.1007/978-3-642-23974-8_43

Michael Neff, Yingying Wang, Rob Abbott, and Marilyn Walker. 2010. Evaluating the effect of gesture and language on personality perception in conversational agents. In *Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.), Lecture Notes in Computer Science, Vol. 6356. Springer Berlin Heidelberg, 222–235. DOI:http://dx.doi.org/10.1007/978-3-642-15892-6_24

Jay F. Nunamaker, Douglas C. Derrick, Aaron C. Elkins, Judee K. Burgoon, and Mark W. Patton. 2011. Embodied conversational agent-based kiosk for automated interviewing. *J. Manag. Inf. Syst.* 28, 1 (2011), 17–48. DOI:http://dx.doi.org/10.2753/MIS0742-1222280102

B. Okun and R. Kantrowitz. 2014. *Effective Helping: Interviewing and Counseling Techniques* (6th ed.). Cengage Learning.

Astrid M. Von Der Pütten, Nicole C. Krämer, and Jonathan Gratch. 2010. How our personality shapes our interactions with virtual characters—Implications for research and development. In *Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.), Lecture Notes in Computer Science, Vol. 6356. Springer Berlin Heidelberg, 208–221. DOI:http://dx.doi.org/10.1007/978-3-642-15892-6_23

Brian Ravenet, Magalie Ochs, and Catherine Pelachaud. 2013. From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes. In *Intelligent Virtual Agents*, Ruth Aylett, Brigitte Krenn, Catherine Pelachaud, and Hiroshi Shimodaira (Eds.), Lecture Notes in Computer Science, Vol. 8108. Springer Berlin Heidelberg, 263–274. DOI:http://dx.doi.org/10.1007/978-3-642-40415-3_23

Jingjing Ren, Daniel Schulman, Brian Jack, and Timothy W. Bickmore. 2014. Supporting longitudinal change in many health behaviors. In *Proceedings of the CHI'14 Extended Abstracts on Human Factors in Computing Systems (CHI EA'14)*. ACM, New York, NY, 1657–1662. DOI:http://dx.doi.org/10.1145/2559206.2581201

Deborah Richards. 2012. Agent-based museum and tour guides: Applying the state of the art. In *Proceedings of the 8th Australasian Conference on Interactive Entertainment: Playing the System (IE'12)*. ACM, New York, NY, Article 15, 9 pages. DOI:http://dx.doi.org/10.1145/2336727.2336742

Virginia Richmond, James McCroskey, and Mark Hickson. 2008. *Nonverbal Communication in Interpersonal Relations* (6th ed.). Allyn and Bacon. 366 pages.

Ronald E. Riggio and Howard S. Friedman. 1986. Impression formation: The role of expressive behavior. *J. Personal. Soc. Psychol.* 50, 2 (1986), 421–427. DOI:http://dx.doi.org/10.1037//0022-3514.50.2.421

Lazlo Ring, Lin Shi, Kathleen Totzke, and Timothy Bickmore. 2014a. Social support agents for older adults: Longitudinal affective computing in the home. *J. Multimodal User Interfaces* 9, 1 (2014), 79–88. DOI:http://dx.doi.org/10.1007/s12193-014-0157-0

Lazlo Ring, Dina Utami, and Timothy Bickmore. 2014b. The right agent for the job? In *Intelligent Virtual Agents*, Timothy Bickmore, Stacy Marsella, and Candace Sidner (Eds.), Lecture Notes in Computer Science, Vol. 8637. Springer International Publishing, 374–384. DOI:http://dx.doi.org/10.1007/978-3-319-09767-1_49

Albert Rizzo, Bruce Sheffield John, Brad Newman, Josh Williams, Arno Hartholt, Clarke Lethin, and John Galen Buckwalter. 2013. Virtual reality as a tool for delivering PTSD exposure therapy and stress resilience training. *Mil. Behav. Health* 1, 1 (2013), 52–58. DOI:http://dx.doi.org/10.1080/21635781.2012.721064

Susan Robinson, David Traum, Midhun Ittycheriah, and Joe Henderer. 2008. What would you Ask a conversational agent? Observations of human-agent dialogues in a museum setting. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)* (28–30), Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, and Daniel Tapias (Eds.), European Language Resources Association (ELRA), Marrakech, Morocco.

Nicholas Rule and Nalini Ambady. 2010a. First impressions of the face: Predicting success. *Soc. Personal. Psychol. Compass* 4, 8 (2010), 506–516. DOI:http://dx.doi.org/10.1111/j.1751-9004.2010.00282.x

Nicholas O. Rule and Nalini Ambady. 2008. Brief exposures: Male sexual orientation is accurately perceived at 50 ms. *J. Exp. Soc. Psychol.* 44, 4 (2008), 1100–1105. DOI:http://dx.doi.org/10.1016/j.jesp.2007.12.001

Nicholas O. Rule and Nalini Ambady. 2010b. Democrats and republicans can be differentiated from their faces. *PLoS ONE* 5, 1 (2010), e8733. DOI:http://dx.doi.org/10.1371/journal.pone.0008733

Raafat George Saadé and Camille Alexandre Otrakji. 2007. First impressions last a lifetime: Effect of interface type on disorientation and cognitive load. *Comput. Hum. Behav.* 23, 1 (2007), 525–535. DOI:http://dx.doi.org/10.1016/j.chb.2004.10.035

Cody Sandifer. 1997. Time-based behaviors at an interactive science museum: Exploring the differences between weekday/weekend and family/nonfamily visitors. *Sci. Educ.* 81, 6 (1997), 689–701. DOI:http://dx.doi.org/10.1002/(SICI)1098-237X(199711)81:6<689::AID-SCE6>3.0.CO;2-E

Gerard Saucier. 1994. Mini-markers: A brief version of goldberg's unipolar big-five markers. *J. Personal. Assess.* 63, 3 (1994), 506. DOI:http://dx.doi.org/10.1207/s15327752jpa6303_8

Daniel Schulman, Mayur Sharma, and Timothy Bickmore. 2008. The identification of users by relational agents. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems—Volume 1 (AAMAS'08)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 105–111. http://dl.acm.org/citation.cfm?id=1402383.1402403

W. C. Schutz. 1958. *FIRO: A Three-dimensional Theory of Interpersonal Behaviour*. Holt, Rinehart and Winston. 267 pages.

Etienne De Sevin, Sylwia Julia Hyniewska, and Catherine Pelachaud. 2010. Influence of personality traits on backchannel selection. In *Intelligent Virtual Agents*, Jan Allbeck, Norman Badler, Timothy Bickmore, Catherine Pelachaud, and Alla Safonova (Eds.), Lecture Notes in Computer Science, Vol. 6356. Springer Berlin Heidelberg, 187–193. DOI:http://dx.doi.org/10.1007/978-3-642-15892-6_20

Deng Shengli, Liu Yong, Li Hongxiu, and Hu Feng. 2013. How does personality matter? An investigation of the impact of extraversion on individuals' SNS use. *Cyberpsychol., Behav., Soc. Netw.* 16, 8 (2013), 575–581. DOI:http://dx.doi.org/10.1089/cyber.2012.0383

Oliviero Stock and Massimo Zancanaro. 2007. *PEACH—Intelligent Interfaces for Museum Visits (Cognitive Technologies)*. Springer-Verlag New York, Inc., Secaucus, NJ.

Michael Sunnafrank. 1986. Predicted outcome value during initial interactions a reformulation of uncertainty reduction theory. *Hum. Commun. Res.* 13, 1 (1986), 3–33. DOI:http://dx.doi.org/10.1111/j.1468-2958.1986.tb00092.x

Michael Sunnafrank and Artemio Ramirez. 2004. At first sight: Persistent relational effects of get-acquainted conversations. *J. Soc. Personal Relationsh.* 21, 3 (2004), 361–379. DOI:http://dx.doi.org/10.1177/0265407504042837

J. Svennevig. 1999. *Getting Acquainted in Conversation*. John Benjamins, Philadephia, PA, USA.

Kazuki Takashima, Yasuko Omori, Yoshiharu Yoshimoto, Yuich Itoh, Yoshifumi Kitamura, and Fumio Kishino. 2008. Effects of avatar's blinking animation on person impressions. In *Proceedings of Graphics Interface 2008 (GI'08)*. Canadian Information Processing Society, Toronto, Ontario, Canada, 169–176. DOI:http://dx.doi.org/10.1145/1375714.1375744

Adriana Tapus, Cristian Ţăpuş, and MajaJ Matarić. 2008. User'robot personality matching and assistive robot behavior adaptation for post-stroke rehabilitation therapy. *Intell. Serv. Robot.* 1, 2 (2008), 169–183. DOI:http://dx.doi.org/10.1007/s11370-008-0017-4

David R. Traum, Priti Aggarwal, Ron Artstein, Susan Foutz, Jillian Gerten, Athanasios Katsamanis, Anton Leuski, Dan Noren, and William R. Swartout. 2012. Ada and grace: Direct interaction with museum visitors. In *Intelligent Virtual Agents*, Yukiko Nakano, Michael Neff, Ana Paiva, and Marilyn Walker (Eds.), Lecture Notes in Computer Science, Vol. 7502. Springer Berlin Heidelberg, 245–251. DOI:http://dx.doi.org/10.1007/978-3-642-33197-8_25

Astrid M. von der Pütten, Nicole C. Krämer, Jonathan Gratch, and Sin-Hwa Kang. 2010. "It doesn't matter what you are!" explaining social effects of agents and avatars. *Comput. Hum. Behav.* 26, 6 (2010), 1641–1650. DOI:http://dx.doi.org/10.1016/j.chb.2010.06.012 Online Interactivity: Role of Technology in Behavior Change.

M.L. Walters, M.A. Oskoei, D.S. Syrdal, and K. Dautenhahn. 2011. A long-term human-robot proxemic study. In *Proceedings of the RO-MAN, 2011 IEEE*. 137–142. DOI:http://dx.doi.org/10.1109/ROMAN.2011.6005274

Ning Wang and Jonathan Gratch. 2009. Can virtual human build rapport and promote learning?. In *Proceedings of the 2009 Conference on Artificial Intelligence in Education*. IOS Press, Amsterdam, The Netherlands, 737–739. DOI:http://dx.doi.org/10.3233/978-1-60750-028-5-737

A. Watson, T. Bickmore, A. Cange, A. Kulshreshtha, and J. Kvedar. 2012. An internet-based virtual coach to promote physical activity adherence in overweight adults: Randomized controlled trial. *J. Med. Internet Res.* 11, 1 (2012), 293–327. DOI:http://dx.doi.org/10.2196/jmir.1629

Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*. ACM, New York, NY, 143–146. DOI:http://dx.doi.org/10.1145/1978942.1978963

Shuo Zhou, Raghavendra Gali, Michael Paasche-Orlow, and Timothy W. Bickmore. 2014. Afraid to ask: Proactive assistance with healthcare documents using eye tracking. In *Proceedings of the CHI'14 Extended Abstracts on Human Factors in Computing Systems (CHI EA'14)*. ACM, New York, NY, 1669–1674. DOI:http://dx.doi.org/10.1145/2559206.2581218

Katja Zibrek and Rachel McDonnell. 2014. Does render style affect perception of personality in virtual humans?. In *Proceedings of the ACM Symposium on Applied Perception (SAP'14)*. ACM, New York, NY, 111–115. DOI:http://dx.doi.org/10.1145/2628257.2628270

Received December 2014; revised May 2016; accepted May 2016