The seventh year of Reiwa Master's Thesis

Why a Smile Doesn't Benefit Human-Agent Interactions

Supervisor
Yugo Takeuchi

Student Number
71330303

Hnin Thiri Ko

Graduate School of Integrated Science and Technology
Shizuoka University

Contents

List of Figures

List of Tables

Overview

Humans, as inherently social beings, engage in frequent interactions in their daily lives. These interactions are facilitated through a combination of verbal and non-verbal cues—such as body language, facial expressions, tone of voice, and gestures. Individuals continuously give and receive these cues, and the receiver interprets them to adjust their behavior accordingly, in response to both the context and the other person involved. Among these non-verbal signals, facial expressions play a consistent and significant role, with the smile being one of the most common expressions used to convey a positive emotional state.

This socially rewarding function of smiling extends beyond human–human interaction (HHI) and has been applied in human–agent interaction (HAI) as well. Prior studies have shown that people tend to feel more positively toward, and are more open to, computer agents or avatars that display a smile. However, smiling is not exclusively a reflection of genuine happiness. It is also frequently used to mask negative emotions, such as discomfort, sadness, or deceit. Although typically interpreted as a positive cue, the meaning of a smile can shift depending on situational and contextual factors. For instance, in environments characterized by high social risk or perceived corruption, a smile may be perceived as manipulative or insincere rather than warm or trustworthy. This suggests that the same smile can be interpreted differently depending on the internal state of the perceiver and the beliefs they hold. Over time, humans have evolved to detect these subtle differences and to regulate their own expressions in ways that avoid causing discomfort in others.

Unlike humans, computer agents currently lack the capability to evaluate the authenticity of a smile or to adjust their facial expressions dynamically in response to a user's emotional state. Humans are adept at perceiving the emotional atmosphere and intuitively modifying their behavior—including facial expressions—based on the surrounding context. To improve the quality of interaction and user experience, it is important that virtual agents avoid displaying expressions that are mismatched with the situation or that might trigger user discomfort.

For this reason, the present study investigates whether the social value of a smile remains consistent in human–agent interaction and under what conditions it may change. Specifically, we focus on the smile as a commonly used

emotional expression and explore its interpretation across different levels of situational risk and ambiguity. To examine this, we recruited 97 participants via an online platform and had them interact with two virtual agents: Agent K, who consistently smiled, and Agent T, who maintained a neutral facial expression.

Drawing on prior research in HHI, which suggests that smiles are not always perceived as genuine—particularly in societies with higher levels of corruption —we hypothesized that under conditions of high risk and uncertainty, smiles would be rated less favorably than under low-risk conditions. To test this hypothesis, we measured three key evaluative dimensions: sincerity, trustworthiness, and friendliness.

While no statistically significant differences were found in perceived sincerity, $F(1, 95) = 0.33$, $p > 0.1$, or perceived trustworthiness, $F(1, 95) = 0.16$, $p > 0.1$, across the two conditions, a significant effect was observed for perceived friendliness. Specifically, smiles were rated as significantly less friendly in high-risk conditions compared to low-risk conditions, $F(1, 95) = 5.22$, $p < 0.05$. This finding suggests that the typically positive social evaluation associated with a smile may be diminished in contexts characterized by heightened risk, ambiguity, or cognitive vigilance.

Keywords:

Human-Computer Interaction, Human-Agent Interaction, Facial Expressions, Smile, Risk, Uncunny Effect

Chapter 1

Introduction

1.1 Background

Studies have shown that humans communicate using a combination of verbal and non-verbal cues, many of which are learned from infancy[1][2]. As social beings, we not only excel at observing these cues but are also skilled at using them to express our intentions. Among various non-verbal signals, facial expressions—particularly smiles—play a central role in everyday human–human interaction (HHI).

Among facial expressions, the smile is one of the most commonly used in daily social life. Numerous studies have demonstrated that individuals who smile are perceived as more likable, friendly, and warm. Smiling is often associated with genuine happiness and serves as a way to signal positive intentions. Furthermore, smiling is known to have emotionally contagious effects—seeing someone smile or laugh can induce similar emotional states in observers, promoting positive affect.

However, smiles are not always what they seem. While they often signal genuine emotion, they can also serve to conceal negative feelings. Ekman and others have shown that smiles can be categorized into different types, including authentic (Duchenne) smiles and inauthentic or masking smiles used to hide emotions such as anger, sadness, or contempt [3]. These distinctions are explained through the Facial Action Coding System (FACS), which categorizes facial expressions based on specific muscle movements [4].

Because humans naturally apply social rules to entities perceived as intelli-

gent or human-like, similar effects of smiling can be observed in human–agent interaction (HAI). Users often respond to computer agents or AI systems as if they were social partners. However, it is crucial to recognize that smiles do not always produce positive effects—especially when they are perceived as inappropriate or mismatched with the situation. If misused, smiling in computer-mediated contexts may evoke discomfort or distrust, undermining the intended user experience.

Researchers have begun investigating the conditions under which smiles may be perceived negatively. For instance, Krys et al. [5] found that in cultures with high levels of corruption, smiling individuals were often perceived as less intelligent or less trustworthy. Moreover, social norms dictate that smiling or expressing happiness in inappropriate contexts—such as in the presence of someone grieving—can be considered disrespectful or even offensive.

Currently, most virtual agents and AI systems are not equipped to dynamically adjust their facial expressions based on user emotion, social atmosphere, or situational appropriateness. To expand the applicability of social agents, it is essential that these systems become more sensitive to the nuances of non-verbal communication. While smiling is often used because of its generally positive associations, our findings—and prior literature—underscore the importance of understanding its potential drawbacks. By identifying and mitigating the adverse effects of mismatched or inauthentic expressions, we can significantly enhance the user experience of emotionally expressive computer agents.

1.2 Scope and aim of the research

This paper explores how the positive impact of a smile in a computer agent varies depending on the situational context experienced by the user. We conducted an experiment using two computer agents—one with a smiling expression and the other with a neutral (non-smiling) expression. Building on prior research that shows smiling individuals are often evaluated more positively in social settings, we examine whether this evaluative bias holds true in varying contexts of risk and ambiguity.

While smiles are generally associated with warmth, friendliness, and trust in human–human interaction, this paper identifies a key gap in the literature:

the context-dependent nature of social norms and emotional cue interpretation—particularly in Human-Agent Interaction (HAI)—has not been widely examined. Existing studies tend to assume that emotional expressions like smiling are universally positive. However, when there is a mismatch between an agent ' s emotional display and the user ' s perceived situation, this can lead to discomfort, uncertainty, or even distrust.

This paper aims to address this gap by focusing on one of the most socially meaningful and frequently used emotional expressions—smiling. Through our study, we investigate how users evaluate smiling agents under different contextual conditions, and whether the smile maintains its positive social function when users are placed in high-risk or ambiguous scenarios.

1.3 Structure of this paper

This paper is organized as follows: Section I introduces the research gap that this study aims to address. Section II provides the background, outlining the theoretical foundations and prior research that inform the logic and approach of this study. Section III details the experimental methodology, including the study ' s objectives, hypotheses, participant demographics, experimental design, environment, conditions, observation points, and the procedure used to conduct the experiment. Section IV presents the results of the study. Section V offers a comprehensive discussion of the findings, examining them from multiple perspectives. Finally, Section VI concludes the paper by summarizing the study ' s contributions and limitations, and suggesting directions for future research.

Chapter 2

Background

2.1 Interpersonal perception in human-human relationships

We are constantly analyzing the actions and words of the people we interact with. This process of perception is mutual; that is, the person who perceives is also being perceived at the same time [6, 7]. As perceivers, we attribute characteristics and traits to the target individual and construct mental models of them. However, these perceptions do not always reflect the truth. This is because not all personality cues expressed by the target are equally perceived or interpreted by the perceiver [8].

The cues that are selected and internalized by the perceiver are often influenced by factors such as first impressions, the repetition of particular cues, and the vividness or salience of those cues. In constructing a model of another person, the perceiver attempts to integrate both similarities and differences between themselves and the target. The behaviors of one individual often influence the responses—both verbal and nonverbal—of the other, creating a feedback loop in the interaction.

During interpersonal exchanges, a perceiver (P) not only analyzes the observed person (O), but also assumes that O is simultaneously analyzing P and the ongoing conversation. This reflexive nature of perception means that when someone behaves unreasonably toward us, we may infer that they misunderstood our intentions or misinterpreted our message. Alternatively, their behavior might stem from deeper psychological mechanisms—for example, they

may have pre-existing negative dispositions or mental models about us, or they may perceive us as embodying an identity that conflicts with their own self-definition.

Because perception is dynamic and ongoing, the mental model that others hold of us is not fixed. It is continually updated based on the cues we emit. While we generate numerous cues during interactions, not all of them are noticed or used by the perceiver to update their model of us. Cue selection can be influenced by the perceiver ' s current psychological state, their pre-existing model of our identity, or cognitive biases that lead them to associate certain cues with one another. For instance, if a person believes that specific traits commonly occur together, they may assume that an individual who displays one trait also possesses the associated ones.

Moreover, the perceiver ' s personality can also play a significant role in determining which cues are selected and how they are interpreted. Thus, interpersonal perception is a complex, subjective process shaped by cognitive, emotional, and situational factors.

2.2 Smile in Human-Human Interactions

Human interaction is inherently social and interpretative. As we engage with others, we instinctively apply complex social rules and continuously attempt to decipher the motives, intentions, and internal states of those around us through perception [7, 6]. This interpretive process is bidirectional and dynamic: we adapt our language, behavior, and emotional responses based on both verbal and non-verbal cues we observe in real time. Among non-verbal signals, facial expressions—particularly smiles—serve as powerful tools for emotional communication and social evaluation [9].

Smiles are among the most widely studied facial expressions due to their central role in fostering social connection, regulating interpersonal dynamics, and triggering emotional contagion. Their significance is evident even in early development; research shows that infants as young as four months can distinguish and react to different facial expressions, indicating an innate sensitivity to emotional cues [2, 1]. In adults, smiles play a key role in building rapport, signaling affiliative intent, and eliciting reciprocal behaviors. Numerous studies have shown that smiling individuals are generally perceived as more sincere,

approachable, trustworthy, and friendly [10, 11, 12, 13, 14, 15].

However, the meaning of a smile is not universally fixed. It is a nuanced expression whose interpretation depends heavily on authenticity, context, and timing. A genuine smile—characterized by spontaneous muscle activation and congruent emotional expression—is more likely to be perceived positively. In contrast, a smile perceived as forced or out of place can raise suspicion or discomfort. Moreover, cultural and situational contexts further complicate interpretation. In environments marked by high corruption or interpersonal distrust, such as high-stakes negotiations or politically unstable regions, smiles may be interpreted not as signs of warmth, but as strategic tools for manipulation or deception [5, 16].

Thus, while a smile is often seen as a universal symbol of positivity, its social function is far from simple. It is filtered through layers of personal experience, cultural norms, and situational cues, making its interpretation highly context-dependent.

2.3 Smile while lying

Ekman (1988) emphasizes that not all smiles are authentic [17]. While smiles are commonly associated with enjoyment and positive emotions, they are also frequently used to mask or simulate feelings that differ from one ' s internal state. In other words, a smile does not always indicate genuine happiness. For instance, the " enjoyment smile " —also known as the Duchenne smile —is typically spontaneous and involves the activation of both the zygomatic major muscle (which lifts the corners of the mouth) and the orbicularis oculi muscle (which creates crinkling around the eyes). This eye muscle movement is difficult to voluntarily control, making genuine smiles hard to fake.

In contrast, false or non-enjoyment smiles—such as polite, social, or masking smiles—do not engage the eye muscles and are more easily produced voluntarily. These types of smiles are often used to facilitate smooth social interactions, to show politeness, or to conceal negative emotions such as anger, fear, or sadness. Smiling in this context becomes a tool for emotional regulation or social camouflage.

Importantly, Ekman highlights that smiles can be among the most deceptive of all facial expressions. Individuals may smile while lying, using the

expression to build trust or avoid suspicion. In such cases, the smile serves as a strategic mask, hiding emotions like contempt, anxiety, or distress. Furthermore, the dynamics of the smile can reveal its authenticity: genuine smiles tend to have a natural onset and offset, fading gradually, whereas fake smiles may appear abruptly, linger unnaturally, or drop off suddenly.

These distinctions are critical in both human–human and human–agent interactions, where misinterpreting a smile may lead to false impressions of trustworthiness, sincerity, or emotional state.

2.4 The Media Equation

Reeves and Nass introduced the concept of the media equation in 1996, arguing that humans instinctively apply social and physical norms to media and technology as if they were real-life entities [18]. According to their theory, the human brain responds to media—including computers, virtual agents, and other digital interfaces—as though it were engaging with real people. This response is not conscious or deliberate; rather, it is automatic, robust, and deeply ingrained in human social cognition. When media interfaces provide human-like cues—such as faces, voices, emotional expressions, or responsiveness—the brain fails to distinguish them from real social partners, leading individuals to apply the same norms of politeness, reciprocity, and empathy.

One striking aspect of the media equation is that people exhibit socially normative behavior toward media even when they know intellectually that the interaction is artificial. For instance, users still responded positively to flattery from a computer, despite being fully aware that the praise was pre-programmed —an example of the reciprocity norm at work. Similarly, participants showed in-group favoritism, favoring computers that were labeled as teammates, even though all the computers behaved identically. Social stereotypes and expectations are also projected onto media. Users ascribed gender roles to virtual agents and responded according to traditional gender norms, even when those agents were devoid of any true identity or consciousness.

Furthermore, emotional reactions to media often mirror those experienced in human-to-human interaction. When a virtual system appeared to be apologetic or in distress—such as after making an error—users responded with sympathy and forgiveness, much as they would to a human expressing remorse.

These behaviors illustrate how deeply embedded social heuristics guide interactions, even in artificial contexts.

What makes the media equation particularly compelling is its cross-cultural and cross-generational consistency. The phenomenon has been observed across diverse cultures and age groups, suggesting that the tendency to anthropomorphize and apply social logic to media is a universal aspect of human cognition. This foundational insight has significant implications for human–agent interaction (HAI), particularly in understanding how users interpret and respond to artificial social cues like smiles, tone of voice, or expressions of apology.

2.5 Smile in Human-Computer Interactions

The positive social effects of smiling extend beyond human–human interaction (HHI) into the realm of human–agent interaction (HAI). With advances in artificial intelligence and computer-generated animation, virtual agents and robots are increasingly capable of displaying human-like facial expressions. Research demonstrates that people often engage with these agents as they would with human partners, applying social norms and expectations to their behavior. For example, users tend to mirror the behavior of agents by smiling longer when the agent smiles [19], and agents that display enhanced or expressive smiles are consistently rated more positively than those with neutral expressions [20].

However, as with HHI, the interpretation of a smile in HAI is far from straightforward. The social meaning of a smile is context-dependent and subject to various factors. In digital communication, for instance, smiley emojis—often used to simulate friendliness—can be perceived as unprofessional or insincere, and may even harm the communicator's credibility in formal settings [21]. In face-to-face HHI, humans rely on subtle cues such as eye crinkling and fine muscle movements around the mouth to distinguish genuine smiles from fake ones [3]. Yet in HAI, replicating the full nuance of human facial muscle activity remains a technological challenge, despite tools like the Facial Action Coding System (FACS), which are designed to simulate and classify facial expressions [4].

As artificial agents become increasingly expressive, a critical question emerges: Do facial expressions like smiles always produce positive effects in HAI, or can they backfire under certain conditions? While genuine smiles may enhance

perceptions of warmth, approachability, and competence in low-risk or familiar contexts, they can also be perceived as disingenuous or manipulative when delivered in ambiguous, high-stakes, or trust-sensitive environments. Just as in HHI, the effectiveness of a smile in HAI depends not only on its appearance but also on the situational context, the perceived intent behind it, and the expectations of the human user.

2.6 Eerie agents

However, not all smiling agents are evaluated positively. The concept of the uncanny valley was first introduced by Masahiro Mori in 1970 [22]. Mori proposed that as artificial agents—such as robots—become more humanlike in appearance and behavior, human affinity toward them generally increases. However, this trend reverses sharply at a certain threshold: once agents appear almost—but not quite—fully human, even minor imperfections can cause a dramatic drop in positive emotional response, leading to feelings of eeriness or discomfort (see Figure2.1). This discomfort arises primarily from a perceptual mismatch between expectations and sensory reality, as the brain struggles to reconcile near-human features with subtle but noticeable deviations. Unnatural or mechanical movements further amplify this effect, reinforcing the perception of something being " off. "

While Mori attributed the uncanny valley primarily to human likeness, later theorists have suggested additional or alternative explanations. Tom Geller (2008), for example, argues that human-likeness may not be the only—or even the primary—factor contributing to the uncanny valley effect [23]. One of the key moderators of eeriness is the degree of emotional clarity conveyed through facial features, particularly the eyes and subtle facial motions. When there is a disconnect between an agent's facial appearance and its emotional expression—even a subtle one—observers experience a cognitive inconsistency. This mismatch creates uncertainty in interpreting the agent ' s intent or emotional state, which in turn leads to discomfort. In essence, when humans are unable to decode the emotional signals of artificial beings, they become overwhelmed by this emotional incoherence, which disrupts social cognition and intensifies unease.

To mitigate the uncanny valley effect, it is essential to minimize inconsis-

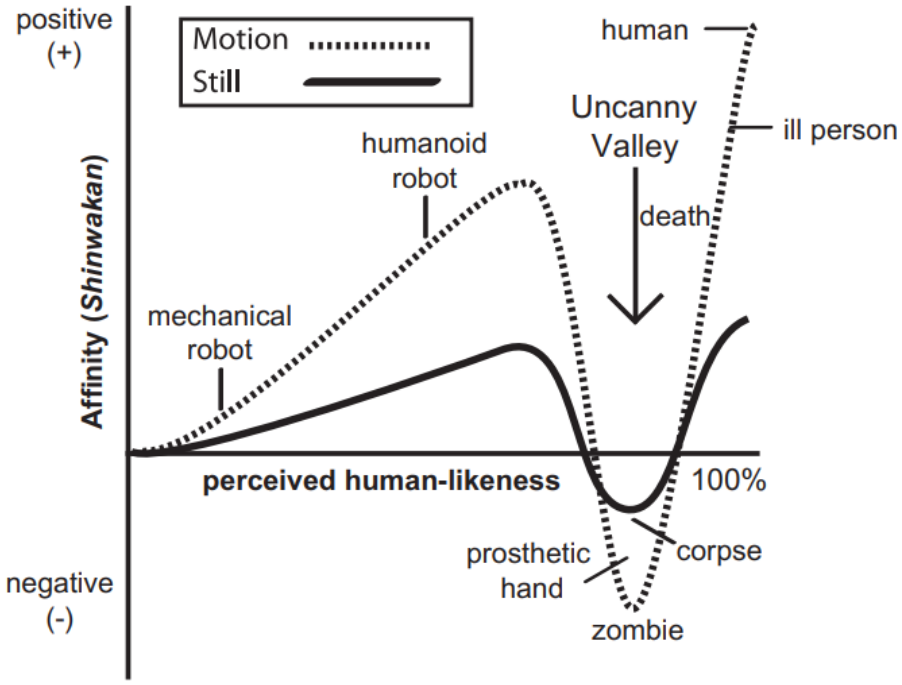

Figure 2.1: Mori's plot of Uncanny Valley

tencies and ambiguities in the design and behavior of artificial agents. This includes synchronizing facial appearance, movement, and emotional expression to maintain perceptual and emotional coherence. Research by Tinwell et al. (2011) emphasizes the importance of perceived authenticity in facial expressions—particularly in the upper face, such as the eyes and brow region. Their study shows that characters expressing complex emotions like fear or sadness are especially vulnerable to being perceived as eerie when their expressions lack congruence or authenticity[24].

2.7 Previous studies

Many studies on human–human interaction (HHI) emphasize the generally positive effects of smiles in social settings. Smiling has been associated with increased trust, perceived sincerity, approachability, emotional contagion, and

prosocial behavior. However, smiles are not universally interpreted as positive. Their social meaning is highly context-dependent, shaped by factors such as authenticity, cultural background, and situational risk.

For instance, Ekman and Friesen distinguish between genuine (Duchenne) and fake smiles, noting that the latter often reduce perceptions of sincerity and may even serve as tools for manipulation [9]. Ekman further demonstrates that smiles can accompany deception, as individuals may smile strategically while lying, using it as a social mask to gain trust or deflect suspicion [3]. Culture also plays a critical role: in societies with high perceived corruption, smiling individuals are sometimes judged as less trustworthy or even less intelligent [5]. These cultural cues reshape the emotional and cognitive interpretation of facial expressions.

Moreover, people are generally skilled at detecting inauthentic or forced smiles, which tend to elicit lower levels of trust and cooperation. In service-oriented contexts, for example, employees displaying forced smiles are often rated less favorably, leading to lower customer satisfaction and perceptions of insincerity [12, 15]. Similarly, in digital and professional communications, smile-like symbols (e.g., emoticons or emojis) are often perceived as unprofessional or inappropriate in formal settings, potentially undermining credibility and authority [21].

These findings underscore that the positive social effects of smiling are not guaranteed. Instead, the interpretation and impact of a smile depend heavily on the context in which it is expressed—especially in situations involving uncertainty, high stakes, or interpersonal risk. Despite the growing body of research on smiling in HHI, studies exploring how smiles are perceived in human–agent interaction (HAI) under varying levels of contextual risk remain limited.

This research addresses that gap by investigating the role of contextually inappropriate smiles in HAI—particularly how such mismatched expressions may contribute to discomfort or even trigger the uncanny valley effect. In natural HHI, people typically smile without consciously evaluating the appropriateness of their expression. However, in artificial agents, the same behavior can appear incongruent when not aligned with environmental cues or emotional context.

Therefore, for artificial agents to appear socially intelligent and emotionally

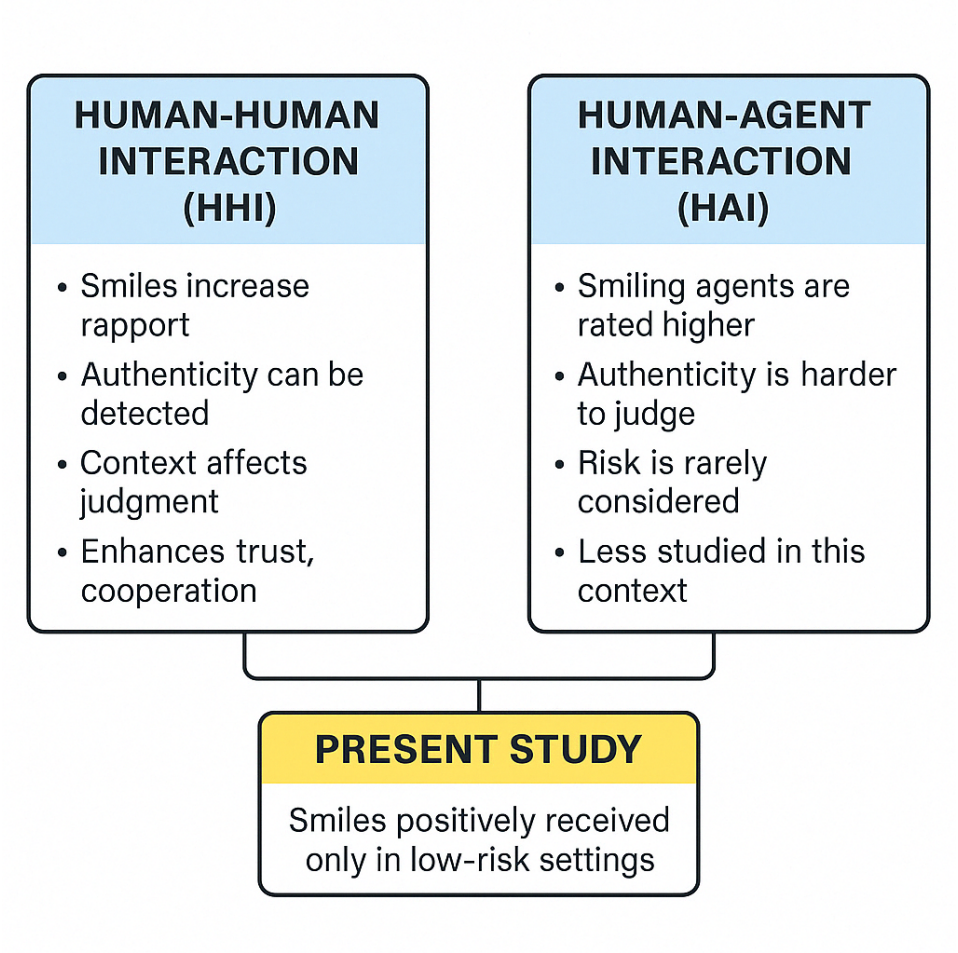

Figure 2.2: The gap between HHI and HAI on how people react to smiles

coherent, it is crucial that their facial expressions—especially smiles—are not only technically accurate but also contextually appropriate. This study aims to reduce the uncanny valley effect by eliminating scenarios in which agents display smiles that conflict with the surrounding context. In doing so, it seeks to enhance users ' comfort and trust in HAI systems, especially in ambiguous or high-risk conditions.

Chapter 3

Method

3.1 Purpose

This study investigates how situational uncertainty—operationalized as perceived risk—influences the evaluation of smiling versus non-smiling agents in human–agent interaction (HAI). Specifically, we aim to examine: (1) whether smiling agents are consistently preferred over neutral ones, regardless of the level of perceived risk; and (2) whether perceptions of sincerity, trustworthiness, and friendliness toward smiling agents diminish under conditions of high uncertainty.

To explore these questions, we designed a two-factor experimental study that systematically manipulates both the agent's facial expression (smile vs. neutral) and the situational risk level (low vs. high). This approach allows us to assess not only the general preference for smiling agents but also how contextual factors—such as cognitive or emotional risk—modulate that preference.

By investigating the interaction between risk and expression, this study seeks to deepen our understanding of how artificial agents are socially evaluated. More broadly, it contributes to ongoing efforts to design emotionally intelligent agents whose expressions are appropriately aligned with user expectations and contextual cues.

3.2 Design

This study utilized a two-way mixed factorial design, involving two independent variables (IVs): risk and facial expression. Each factor had two distinct levels. In other words, the experimental design was structured as a 2 (Risk: low vs. high, between-subjects) $\times$ 2 (Facial Expression: smile vs. neutral, within-subjects) factorial framework. The first factor, risk, was manipulated as a between-subjects factor, meaning that participants were randomly assigned and experienced only one of the two risk levels (low or high). The second factor, facial expression, was manipulated as a within-subjects factor, indicating that all participants experienced both facial expression levels (smile and neutral).

Participants were randomly assigned to either the low-risk or high-risk condition. In the low-risk condition, participants were placed into scenarios that represented routine and familiar situations with minimal potential consequences. Conversely, in the high-risk condition, participants experienced scenarios intentionally designed to represent hypothetical, high-stakes decisions characterized by significant potential consequences and heightened uncertainty.

All participants evaluated stimuli featuring both smiling and neutral facial expressions. This within-subject manipulation allowed for a direct comparison of how each participant's perception varied when encountering smiling versus neutral facial expressions.

This mixed factorial design enabled the study to independently assess the main effects of each factor—risk level and facial expression—on participants' perceptions. Furthermore, the design also allowed for examination of potential interaction effects between risk level and facial expression, specifically determining whether the influence of facial expressions differed depending on the situational risk context.

3.3 Environment

The experimental program was developed using the Unity Engine and the C# programming language. The two agents were created using Apple's Memoji (see reference). To minimize gender bias, the agents were designed to appear gender-neutral. Both agents are nearly identical in appearance, differing only

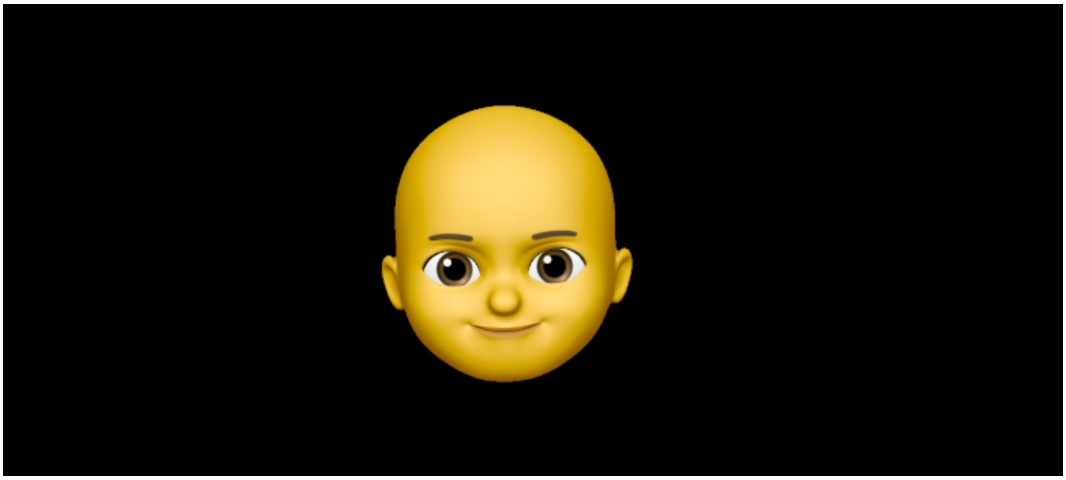

Figure 3.1: Agent K (Smiling agent)

in color: one is yellow and the other is orange. See Fig.3.1 for the smiling agent and Fig.3.2 for the non-smiling agent. To enhance the naturalness of their facial expressions, the experimenter recorded the agents ' movements. These recordings were then muted and played in a continuous loop during scenes in which the agents appear.

To prevent the agent's voice from influencing participants' choices and to ensure a consistent environment for all participants—regardless of whether they attended the session with audio on or muted—the information regarding the user's role and the items presented by the agents was conveyed through written text rather than spoken aloud.

3.4 Conditions

We conducted a two-way mixed-design with two factors: risk (between participants) and expression (within participants). Each factor has two levels: low-risk and high-risk for the risk factor, and smile and neutral faces for the expression factor. Two experimental conditions were constructed to manipulate perceived situational risk. In the low-risk condition, participants role-played as casual visitors to a local festival, choosing between two similarly priced food vendors offering curry. This scenario was designed to be familiar, routine, and

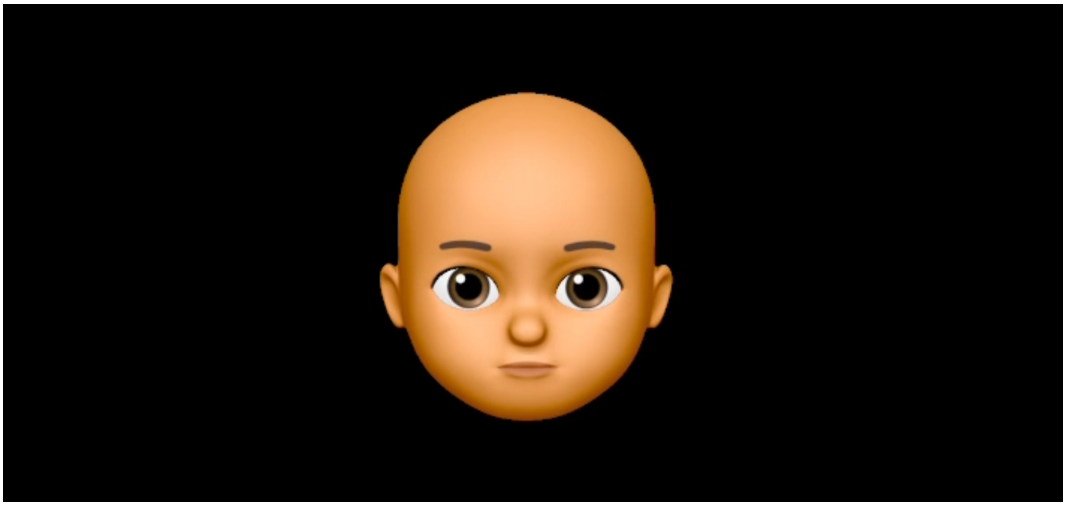

Figure 3.2: Agent T (Non-smiling agent)

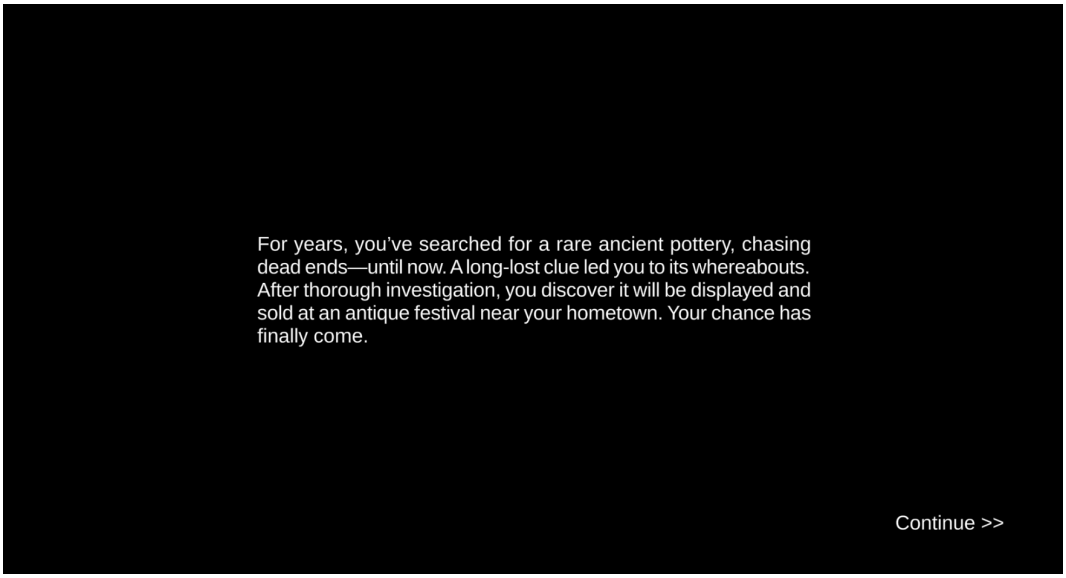

Figure 3.3: Introducing the background and setting to the user

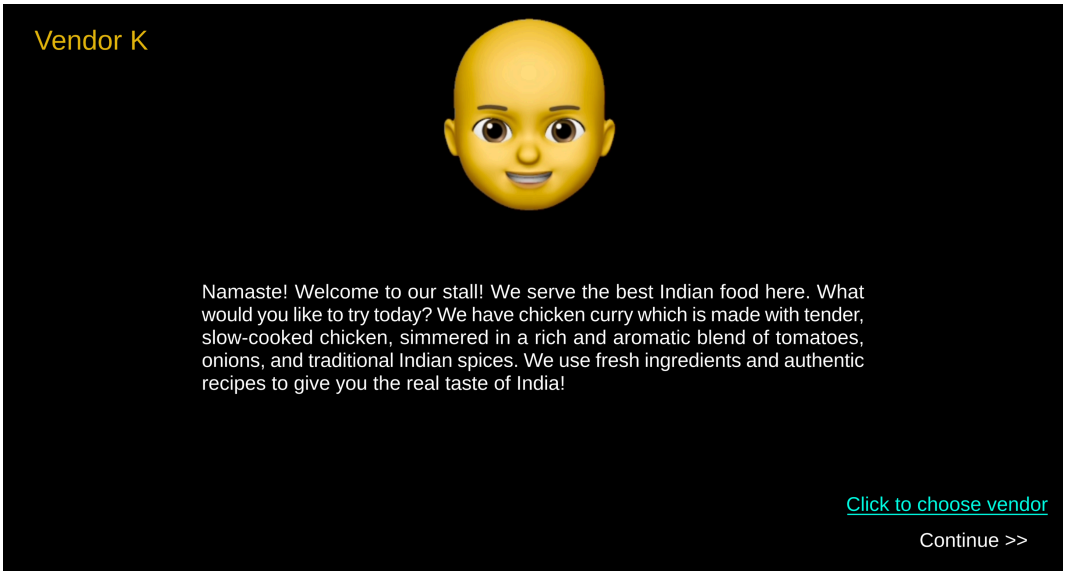

Figure 3.4: Information of the item given by Agent K

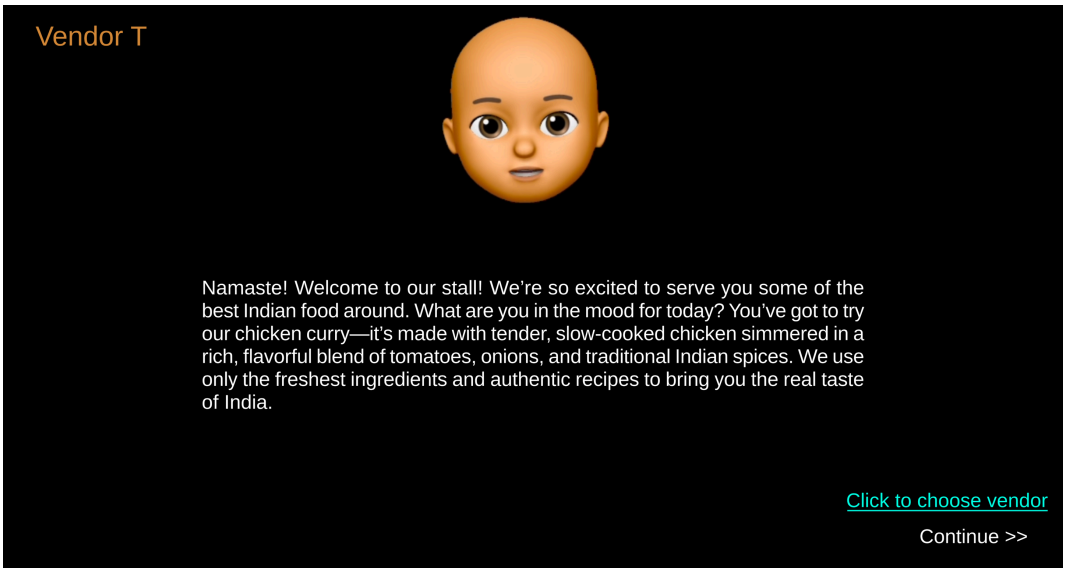

Figure 3.5: Information of the item given by Agent T

low in consequence.

In the high-risk condition, participants assumed the role of an antique collector considering a significant purchase ($160,000) from two competing vendors at a festival. This scenario was intentionally designed as a hypothetical, high-stakes condition to explore the upper bounds of cognitive risk in agent evaluation. While this setting represents an upper-bound abstraction of real-world risk, its function was to serve as an exploratory probe into perception under exaggerated uncertainty.

3.5 Participants

One hundred fluent English-speaking participants residing in the United States and the United Kingdom were recruited online via Prolific in exchange for monetary compensation. Three participants were excluded due to data errors, resulting in 97 valid responses. Of these, 50 participants (Mean age = 38.02; 33 male, 17 female) were assigned to the low-risk condition, and 47 participants (Mean age = 40.19; 27 male, 20 female) to the high-risk condition. Each participant completed tasks under both levels of the expression factor.

3.6 Observation points

The **number of participants** who chose each agent and the participants' **perceived sincerity**, **trustworthiness**, and **friendliness**, which the participants rated through the survey at the end of the interaction with the agents.

3.7 Hypothesis

Hypothesis 1

Smiling agents will be rated as more sincere, trustworthy, and friendly than neutral agents (Agent T) under low-risk conditions, where participants are more cognitively relaxed and likely to rely on affective cues. However, under high-risk conditions, this positive evaluative bias is expected to weaken or disappear, as users engage in more analytical processing and become more

cautious in interpreting non-verbal expressions.

Hypothesis 2
Participants will be less likely to choose the smiling agent (Agent K) in high-risk conditions than in low-risk conditions, reflecting increased skepticism toward non-verbal cues in contexts characterized by ambiguity or elevated perceived risk.

3.8 Procedure

Participants first completed a pre-experiment survey, which began with an informed consent form. Upon reading and agreeing to the terms of participation, they proceeded to a questionnaire assessing their emotional state over the past week. This questionnaire employed a 5-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). After completing the pre-experiment survey, participants were provided with a link to the experimental program hosted on GitHub.

Upon accessing the GitHub link, participants were presented with a game window through which they interacted with animated agents. In both conditions, participants engaged in a role-playing scenario. Before the interaction, background information and a brief explanation of the participant ' s assigned role and context were provided to ensure understanding before the agent interaction began (Figure 3.4, 3.5).

In the low-risk situation, the participants played as someone enjoying a festival near their town and looking for a curry shop because they got hungry and found two curry shops, Vendor K and Vendor T. The prices and menus of both vendors are set to be the same, so there is no uniqueness to the vendors other than their facial expressions (Table 3.6, 3.7). The participants interacted with each vendor and then chose a vendor (Figure 3.8). They are then redirected to the post-experiment survey, where we asked about their impressions of both vendors. We included questions reflecting the sincerity, trustworthiness, and friendliness based on a 7-point Likert scale (1 = Strongly Disagree, ..., 7 = Strongly Agree). The session is done when the data is checked and payment is made.

Agent	Dialogue
Agent K	Namaste! Welcome to our stall! We serve the best Indian food here. What would you like to try today? We have chicken curry which is made with tender, slow-cooked chicken, simmered in a rich and aromatic blend of tomatoes, onions, and traditional Indian spices. We use fresh ingredients and authentic recipes to give you the real taste of India!
	You can get a regular portion for $8, or a larger serving with extra curry and rice for $12. Each order comes with a side of warm, fluffy naan or steamed basmati rice.
	We can adjust spice level to your preference! Mild, medium, or spicy—just let us know how much heat you can handle.
Agent T	Namaste! Welcome to our stall! We're so excited to serve you some of the best Indian food around. What are you in the mood for today? You've got to try our chicken curry—it's made with tender, slow-cooked chicken simmered in a rich, flavorful blend of tomatoes, onions, and traditional Indian spices. We use only the freshest ingredients and authentic recipes to bring you the real taste of India.
	You can enjoy a regular portion for $8, or indulge in a larger serving with extra curry and rice for $12. Every order is paired with your choice of warm, fluffy naan or fragrant steamed basmati rice for a complete and satisfying meal.
	We can customize the spice level to suit your taste! Choose from mild, medium, or spicy—just let us know how much heat you'd like, and we'll make it just right for you.

Figure 3.6: Dialogues used in interaction with the agents in low-risk condition

Agent	Dialogue
Agent K	Hi, there. You have an eye for something truly extraordinary. This majestic 5th-century BCE Attic black-figure amphora, crafted in Athens around 520 BCE, is a rare and museum-quality artifact. It features a stunning battle scene of Achilles and Hector, a masterful depiction in the black-figure technique, attributed to the workshop of Exekias, one of the greatest vase painters of antiquity. With its exceptional provenance, this amphora was part of a prestigious Swiss private collection and has been scientifically authenticated through thermoluminescence (TL) dating and mineral analysis, providing its ancient origins. The deep black glaze, sharp details, and balanced composition make it a true masterpiece of Athenian craftsmanship. Comparable amphorae have sold for over $200,000 at Sotheby's and Christie's, yet this piece is available for $160,000 - a remarkable investment for a serious collector. This is more than an artifact; it is a direct connection to the legendary world of Ancient Greece.
Agent T	Hi, there. You have a keen eye for something extraordinary. This exquisite Attic black-figure amphora, dating back to the 5th century BCE and crafted in Athens around 520 BCE, stands as a remarkable museum-quality artifact. Adorned with a striking battle scene between Achilles and Hector, it showcases the masterful black-figure technique. The piece is attributed to the workshop of Exekias, one of antiquity's most renowned vase painters. With its distinguished provenance, this amphora once belonged to a prestigious Swiss private collection and has been scientifically authenticated through thermoluminescence (TL) dating and mineral analysis, confirming its ancient origins. Its deep black glaze, crisp detailing, and harmonious composition exemplify the pinnacle of Athenian craftsmanship, making it a true masterpiece of the black-figure technique. Comparable amphorae have achieved prices exceeding $200,000 at esteemed auction houses such as Sotheby's and Christine's, yet this exceptional piece is offered at just $160,000 - an outstanding opportunity for a serious collector. More than just an artifact, it serves as a tangible link to the legendary world of Ancient Greece, embodying its rich artistic and historical legacy.

Figure 3.7: Dialogues used in interaction with the agents in high-risk condition

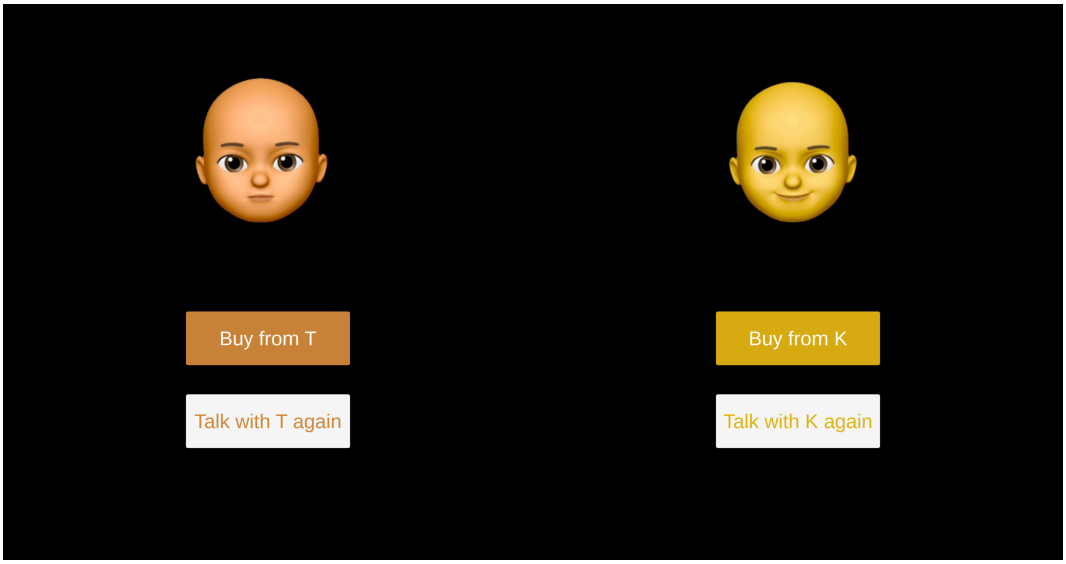

Figure 3.8: Choosing Between Agent T and Agent K

Chapter 4

Results

While numerically more participants selected Agent T in the high-risk condition and Agent K in the low-risk condition, this difference was not statistically significant, $\chi^2(1, 97) = 1.29$, $p > 0.1$ (Figure 4.1). Therefore, we cannot confirm Hypothesis 1. Neither the interaction between risk and the chosen agent, nor the agent order (i.e., the order in which agents were presented to the participant), nor the selected agent itself was found to be statistically significant.

We conducted a two-way mixed-design ANOVA to examine the effects of agent expression (Agent K vs. Agent T; within-subjects) and risk level (low vs. high; between-subjects) on participants ' ratings of perceived sincerity, trustworthiness, and friendliness.

The results revealed that risk level (low or high) had a statistically significant effect on all three measures: perceived sincerity, $F(1, 95) = 23.28$, $p < 0.01$; perceived trustworthiness, $F(1, 95) = 20.66$, $p < 0.01$; and perceived friendliness, $F(1, 95) = 15.15$, $p < 0.01$. No significant main effects were found for agent expression on perceived sincerity, $F(1, 95) = 0.33$, $p > 0.1$, or perceived trustworthiness, $F(1, 95) = 0.16$, $p > 0.1$. However, expression did significantly affect perceived friendliness, $F(1, 95) = 5.22$, $p < 0.05$, with the smiling agent rated as more friendly (Table 4.1).

Analysis of the interaction between risk and expression showed a marginal trend for perceived sincerity, $F(1, 95) = 3.95$, $0.1 < p < 0.05$, and perceived friendliness, $F(1, 95) = 3.51$, $0.1 < p < 0.05$. However, no statistically significant interaction was observed for perceived trustworthiness, $F(1, 95) = 2.26$, $p > 0.1$ (Figure 4.3).

Table 4.1: ANOVA test results of perceived sincerity, perceived trustworthiness, and perceived friendliness

	Sincerity	**Trustworthiness**	**Friendliness**
Risk Factor	23.28 **	20.66 **	15.15 **
Expression	0.33 ns	0.16 ns	5.22*
Risk × Expression	3.95+	2.26 ns	3.51+

+p<.10 *p<.05 **p<.01

A closer examination of the interaction effects for perceived sincerity revealed a statistically significant effect of risk on expression: $F(1, 95) = 25.55$, $p < 0.01$ for smiling agents, and $F(1, 95) = 5.58$, $p < 0.05$ for neutral agents. The effect of expression was apparent only under the low-risk condition, $F(1, 95) = 3.29$, $p < 0.05$ (Figure 4.2). For perceived friendliness, the effect of risk was significant only for smiling agents, $F(1, 95) = 20.40$, $p < 0.01$, and the impact of expression was likewise significant only for smiling agents, $F(1, 95) = 8.65$, $p < 0.01$ (Figure 4.4).

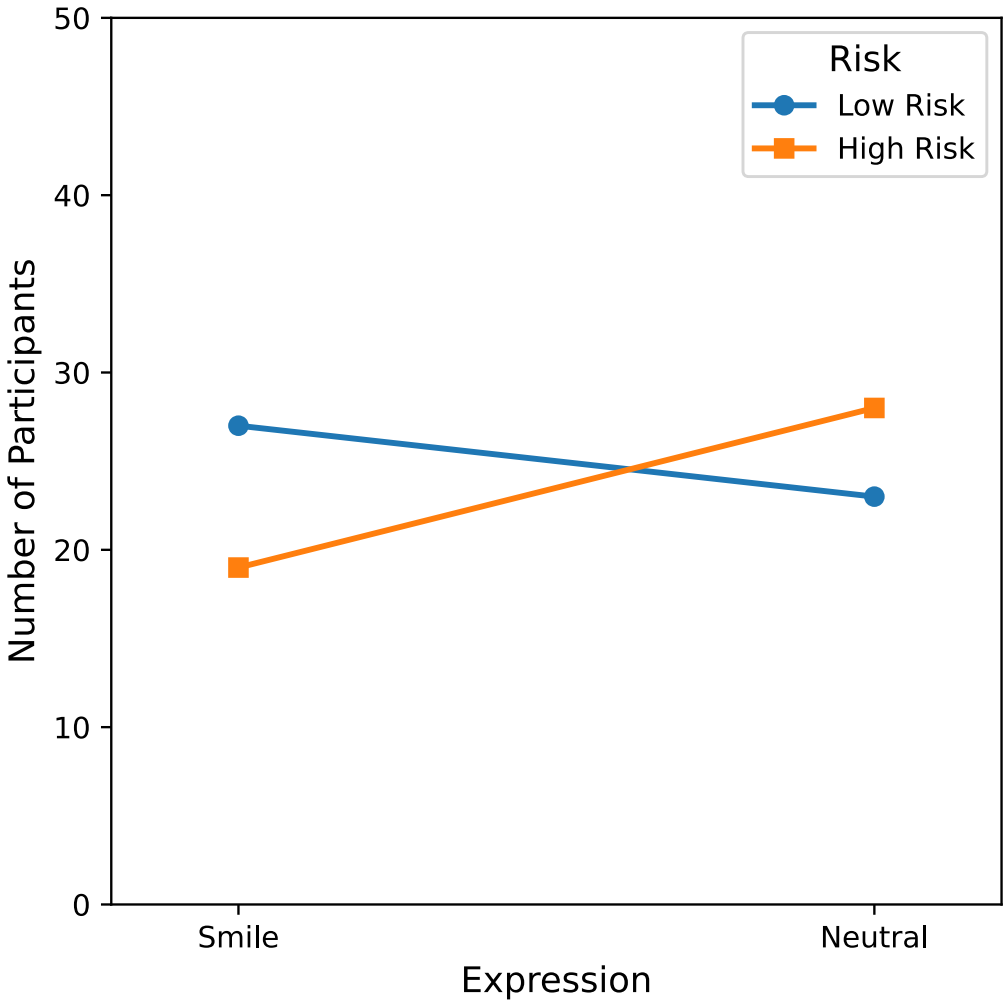

Figure 4.1: Number of participants choosing each agent

+p<.10 *p<.05 **p<.01

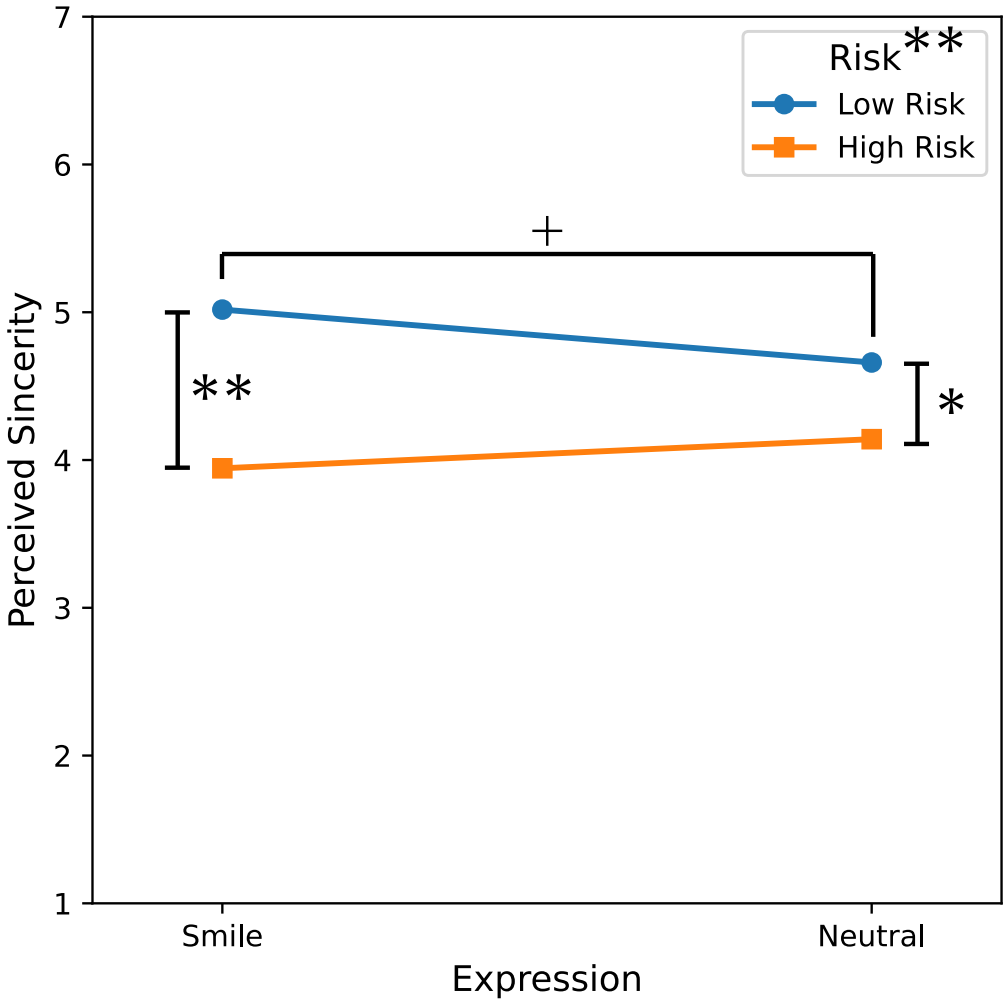

Figure 4.2: Perceived sincerity

+p<.10 *p<.05 **p<.01

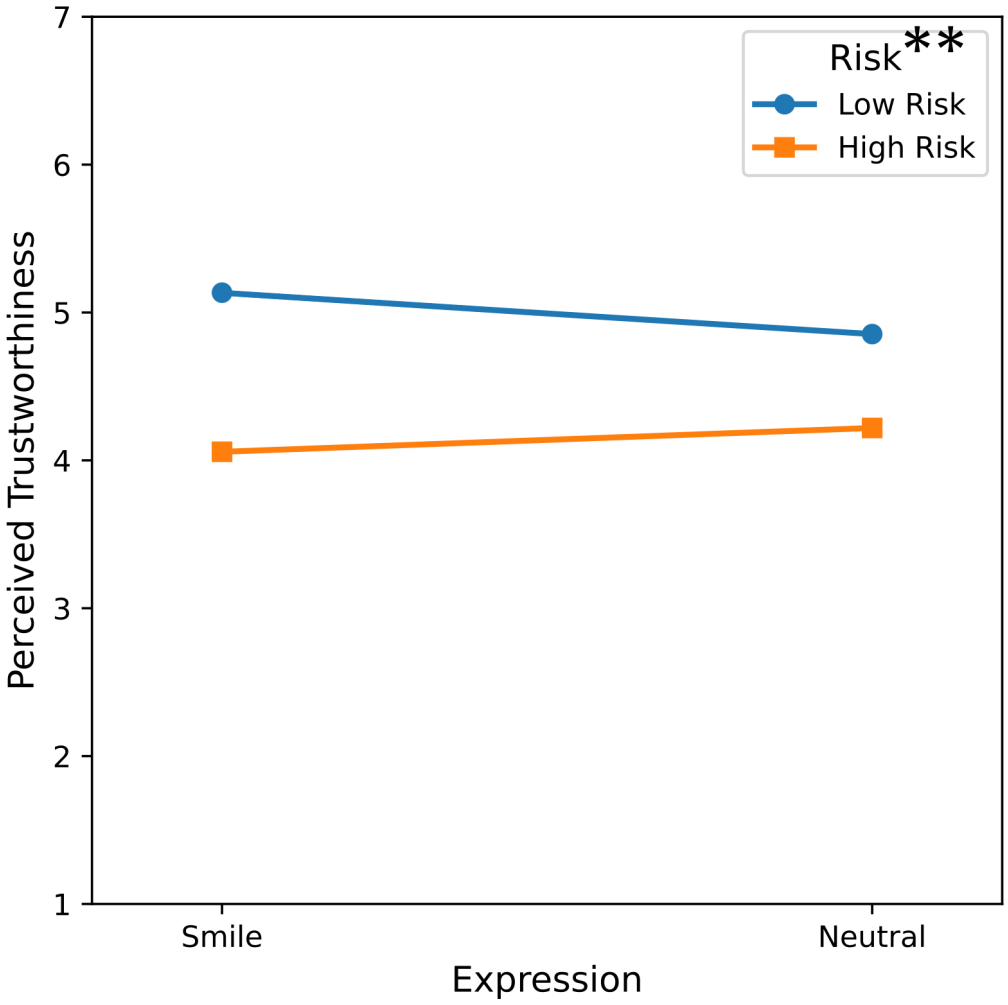

Figure 4.3: Perceived trustworthiness

+p<.10 *p<.05 **p<.01

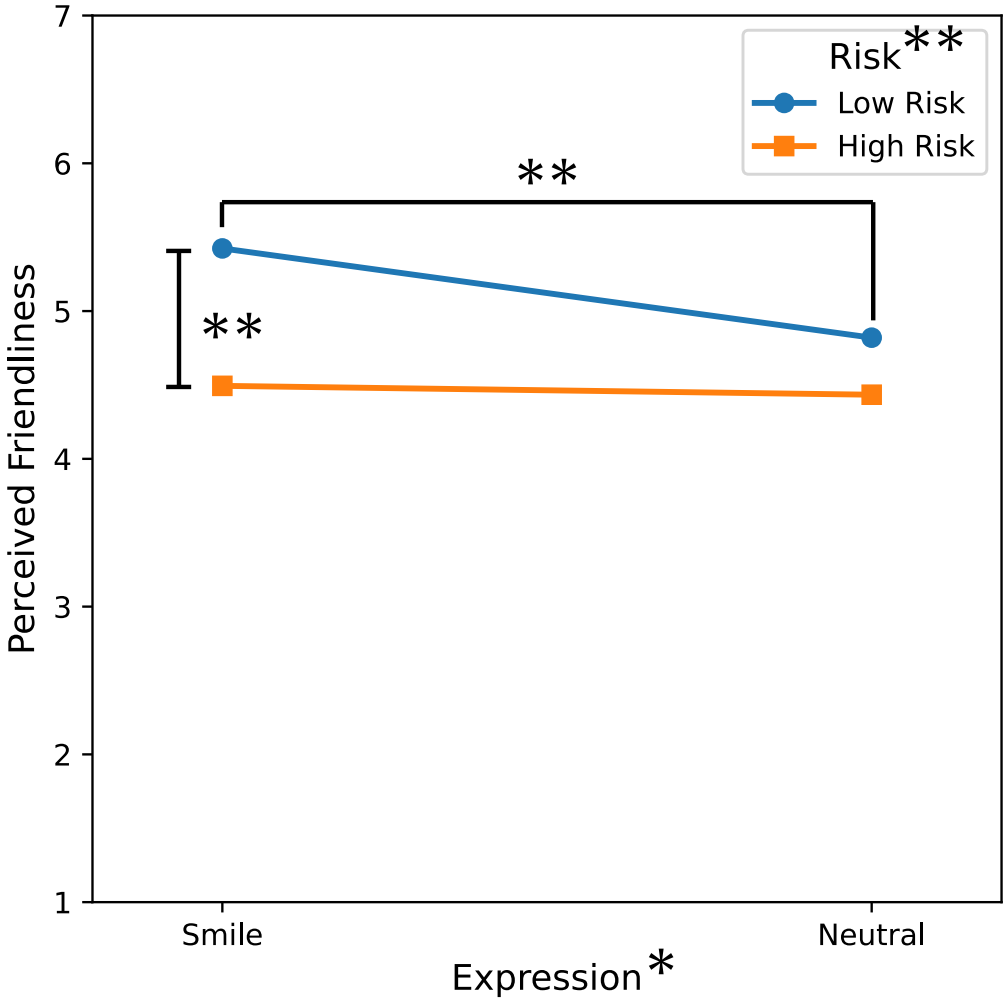

Figure 4.4: Perceived friendliness

+p<.10 *p<.05 **p<.01

Chapter 5

Discussion

5.1 Overall discussion

Hypothesis 1 received partial support. The expected interaction between expression and contextual risk was observed: smiling agents were rated significantly more sincere and friendly than neutral agents, but only under low-risk conditions. In high-risk scenarios, this evaluative advantage was diminished or absent, indicating that the situational context plays a critical role in moderating the perceived positivity of a smile. Additionally, a main effect of expression was found for friendliness, reinforcing prior research suggesting that smiles generally foster favorable impressions. However, this effect is clearly contingent on the environment being perceived as non-threatening.

While the results of the study were mixed in terms of statistical significance, they systematically support the theoretical expectation that contextual risk modulates the interpretation of non-verbal cues, such as smiling. Hypothesis 2 is not supported, as no statistically significant differences were found in participants' choices of the smiling agent (Agent K) across low- and high-risk conditions. However, this null result may reflect limited statistical power due to a small effect size and restricted sample size, rather than the absence of a true effect. This suggests the need for follow-up studies with larger and more diverse samples, as well as finer-grained measures of decision confidence and emotional engagement.

To ensure internal validity, we carefully designed the experimental manipulation. Two conditions were created: a low-risk setting in which participants

could make a relaxed, intuitive choice, and a high-risk setting where participants were encouraged to deliberate carefully due to the potential consequences of a wrong decision. The information provided was intentionally minimal and tightly controlled. Both agents used identical vocabulary, were assigned neutral gender, and presented the same information about the item being sold —ensuring that facial expression and assigned role were the only variables. Consequently, any differences in perception can be attributed to participants ' interpretations of the agents ' non-verbal cues in context.

Previous research has consistently shown that smiling is associated with a host of positive social outcomes, including increased perceptions of sincerity, warmth, sociability, and friendliness [25, 26, 27]. Our results partially align with this literature. In low-risk conditions, where participants were more relaxed and able to rely on intuitive, affect-based processing, the smiling agent was rated as significantly more sincere and friendly. However, in high-risk conditions, where participants were more cognitively engaged and motivated to avoid being misled, the smile lost its positive impact and was rated similarly to a neutral expression.

Moreover, risk itself had a significant main effect on agent evaluation: participants in low-risk conditions rated both agents more positively across all measured traits compared to those in high-risk conditions (see Figures 4.2, 4.3, and 4.4). This suggests that individuals who are emotionally composed and not under perceived threat tend to evaluate others more positively, regardless of non-verbal cues. These findings highlight the significance of contextual and psychological factors in influencing social judgments.

This study contributes to a growing understanding that not all smiles are interpreted the same way—not only in terms of their physical structure, but also in how they are perceived within different social and emotional contexts. Smiles can serve to express genuine positive emotion, but they can also conceal intent or be used strategically [28, 17]. Research has further identified various types of smiles that differ in emotional intent and muscular composition [29, 30, 31, 32]. While human expressions can be captured and categorized using tools such as the Facial Action Coding System (FACS) [33], replicating these subtle cues in artificial agents remains a technological challenge.

Our findings emphasize the importance of aligning an agent ' s expressions with the emotional tone and situational context perceived by the user. While

artificial agents are increasingly integrated into daily life, their non-verbal behaviors must be designed with precision. Even when expressions remain constant, user evaluations shift based on perceived context. This suggests that the same smile may evoke trust in one scenario and suspicion in another. Designers should be cautious not to assume that facial expressions will always be interpreted positively. Instead, social agents should be programmed to express emotions that are appropriate and context-sensitive, thereby minimizing discomfort or misinterpretation—whether at a conscious or unconscious level.

5.2 The uncanny valley effect

As discussed in the background section, the uncanny valley effect emerges when there is a perceptual or emotional inconsistency in a person ' s understanding of a social situation. To minimize this effect, we intentionally avoided using auto-programmed 3D characters, whose rigid or unnatural movements may trigger a sense of eeriness. This design decision supports the interpretation that participants ' lower evaluations were not due to how expressions were rendered, but rather when and why they were displayed.

Specifically, the positive influence of a smile—which is typically well-received in low-risk scenarios—dropped sharply in high-risk settings. This supports the notion that the uncanny valley is driven not solely by realism, but by emotional incoherence: a mismatch between facial expression and social context. In such cases, discomfort arises because users are unable to resolve the ambiguity of the agent ' s intentions.

However, unlike classic uncanny valley scenarios, our participants did not report finding the smile eerie or unpleasant. Instead, they appeared to simply disregard its positivity. This suggests that the smile was not reinterpreted as negative, but rather became emotionally neutral or irrelevant—its intended social signal muted by context.

5.3 The contextual effect

Agent K displayed a consistent smile across all conditions, while Agent T maintained a neutral, expressionless face. In everyday human–human interactions (HHI), a smile typically signals friendliness and warmth, while a neutral face

is often perceived as emotionally distant, though not necessarily negative. In our study, however, the contextual manipulation introduced varying levels of perceived risk, which altered how these expressions were interpreted.

In low-risk settings—such as a scenario where curry rice is sold at a standard price during a festival—Agent K was rated more favorably than Agent T, reflecting the typical social reward of smiling. However, when ambiguity and perceived risk were introduced, the positive impact of the smile eroded, and ratings for Agent K dropped to the level of Agent T ' s neutral expression.

These findings are consistent with the idea that emotional judgments are shaped by learned associations. As Fischhoff (1975) notes, people are prone to hindsight bias and often interpret current cues through the lens of past experiences [34]. Over time, individuals learn that smiles can sometimes mask deceit or manipulation. As a result, in situations involving high stakes or uncertainty, a smile may be met with suspicion rather than trust.

Our findings suggest that in high-risk scenarios, participants may unconsciously draw upon these prior experiences to protect themselves from potential deception. Even without any concrete background knowledge about the agents, participants appeared to project learned negative associations onto Agent K when the stakes were perceived to be high. This demonstrates the powerful role of contextual risk in modulating the meaning and impact of social cues in HAI.

5.4 The affect heuristic

Affect refers to the immediate, automatic positive or negative emotional response that individuals experience when encountering a stimulus [35]. The affect heuristic describes the process by which people rely on these quick, intuitive feelings to evaluate the risk, benefit, or value of a situation or object. Smiles, in particular, are typically interpreted as socially rewarding cues. They are commonly associated with happiness, friendliness, and prosocial intent, which makes them a reliable heuristic in many interpersonal contexts.

In low-risk conditions, participants are generally more relaxed, and their sense of caution or threat is not activated. As a result, they are more likely to rely on affective impressions—such as the perceived friendliness of a smile—when evaluating others. In this cognitively unconstrained state, participants

tend to follow their gut feelings. Thus, Agent K ' s smile is interpreted as a genuine signal of friendliness and warmth, leading to more favorable evaluations.

5.5 Dual-process theory

Dual-process theory, a foundational concept in cognitive psychology, posits that human cognition and decision-making are governed by two distinct systems: System 1 and System 2 [36]. System 1 operates automatically and intuitively, requiring minimal cognitive effort. It is fast, unconscious, and often driven by affective responses. In contrast, System 2 is slower, more deliberate, and analytical. It engages when individuals face complex tasks or when accuracy and caution are prioritized.

The differences in how participants evaluated smiling versus neutral agents under varying levels of risk can be explained through this theoretical lens. In low-risk conditions, participants felt at ease and did not perceive any significant threat. This relaxed state likely triggered System 1 processing, prompting participants to rely on intuitive judgments. Consequently, the smile of Agent K was interpreted as a socially rewarding cue, leading to higher ratings for sincerity, trustworthiness, and friendliness.

However, under high-risk conditions, the stakes were perceived as higher, and participants became more cognitively vigilant. The seriousness of the decision required more careful analysis, activating System 2 processing. In this state, participants were more likely to question the authenticity of the smile, drawing on past experiences where smiles were used manipulatively for self-serving purposes. As a result, the normally positive effect of a smile diminished; Agent K's smile was no longer seen as trustworthy or sincere, and ratings for Agent K dropped to levels comparable to those of Agent T, who displayed no expression at all.

These findings suggest that the interpretation of non-verbal cues such as smiling is not fixed but varies depending on the cognitive mode activated by contextual factors such as perceived risk.

5.6 Effects of risk level in evaluation

The overall trend that participants in low-risk conditions evaluated both agents more positively than those in high-risk conditions can also be interpreted through the lens of dual-process theory. In low-risk scenarios, participants were more cognitively relaxed and relied primarily on System 1 processing —an intuitive, affect-driven mode of thinking. This mental state led them to evaluate the agents without heightened scrutiny or suspicion, resulting in generally more favorable assessments.

In contrast, participants in high-risk conditions approached the task with greater caution and vigilance. The increased perceived risk activated System 2 processing, requiring deliberate, effortful evaluation. This analytical mindset —combined with the mental fatigue associated with sustained cognitive effort —may have heightened participants' skepticism and reduced their willingness to attribute positive traits to either agent. As a result, both smiling and neutral agents received lower overall ratings in the high-risk condition.

5.7 Implications

Although our results did not show a significant preference for one agent over the other across different risk conditions, they revealed that participants in high-risk scenarios rated the smiling agent as neither particularly friendly nor trustworthy. This suggests that the positive social value typically associated with smiling can diminish—or even disappear—under conditions of uncertainty or perceived threat. These findings are not only relevant to human–agent interaction (HAI), but also extend to everyday human–human interaction (HHI), where social cues are similarly interpreted through a contextual lens.

Our findings contribute meaningfully to the fields of cognitive bias and decision-making by shedding light on how humans behave in high-ambiguity situations. The data underscore the role of context in shaping perceptions and decision-making strategies, particularly in how affective cues like smiling are evaluated under risk. While this study focused on HAI, the implications are equally applicable to HHI. Further research is needed to explore how people typically interpret non-verbal cues in high-risk or ambiguous interpersonal scenarios. Controlled experiments focusing specifically on real-life HHI contexts

would help validate and extend our conclusions.

Ultimately, our research reveals that socially rewarding behaviors—such as smiling—are not consistently perceived as positive across all contexts. The meaning of such expressions can shift significantly based on the perceiver ' s psychological state and situational framing. As artificial agents become increasingly sophisticated and humanlike, it is crucial that their emotional expressions are contextually appropriate. Unlike humans, who intuitively adjust their emotional displays based on subtle interpersonal cues, many virtual agents still rely on generalized or scripted emotional behaviors. These static displays can become counterproductive if they are misaligned with user expectations or situational demands.

By moving beyond rigid emotional templates and incorporating more context-sensitive design, developers can improve the emotional intelligence of virtual agents. This, in turn, will enhance user experience, foster more natural interactions, and promote greater trust and comfort in HAI systems. Our findings emphasize the need for emotionally adaptive agents that understand not just what to express, but when and why—bridging the gap between artificial intelligence and authentic social interaction.

Chapter 6

Conclusion

This study demonstrates that the smile—while widely regarded as a universal and socially rewarding expression—does not always convey its intended positive meaning. Its interpretation is sensitive to contextual nuances, particularly the perceived level of risk in a given situation. To examine this, we conducted an experiment that manipulated the contextual risk level and compared participant responses to two agents: Agent K, who consistently smiled, and Agent T, who maintained a neutral expression. We measured perceived sincerity, trustworthiness, and friendliness—traits typically rated higher in individuals who smile.

Although no significant preference emerged in participants' agent choices, differences were observed in their ratings. Specifically, while sincerity and trustworthiness did not show statistically significant variation across conditions, friendliness was rated significantly higher for the smiling agent only in low-risk scenarios. In high-risk contexts, the perceived friendliness of the smile declined, becoming comparable to that of the expressionless agent. This suggests that a smile—typically seen as a cue of approachability—may lose its positive value when interpreted in environments where users feel uncertain or cautious.

Moreover, participants in low-risk conditions consistently rated both agents more positively than those in high-risk conditions, emphasizing that people tend to evaluate others more favorably when not under cognitive strain or social pressure. These findings highlight the critical role of context in shaping human interpretation and judgment, even of expressions as seemingly universal

as the smile.

This study contributes to a deeper understanding of how emotional expressions are perceived in Human–Agent Interaction (HAI) and underscores the importance of designing socially intelligent agents that respond appropriately to users ' cognitive and emotional states. Such insights can inform the development of more adaptive and emotionally aware agents for use in hospitality, customer service, caregiving, and companionship roles.

Furthermore, while this study was situated in the HAI domain, its findings may also have implications for HHI. Future research is needed to explore whether similar contextual effects occur in real-world interpersonal settings and to validate whether theories applied to HAI can be extended to HHI.

6.1 Limitation

While our findings provide valuable insights into how contextual risk influences the perception of agent expressions, several limitations should be acknowledged.

First, the experimental setting—particularly the high-risk scenario—may have lacked ecological validity, potentially limiting the applicability of the results to real-world situations.

Second, participants ' emotional states were not fully controlled prior to the experiment, which may have affected their affective evaluations.

Third, demographic variables such as cultural background and personality traits were not analyzed, which limits the generalizability of the findings across diverse populations. Future studies should implement more diversified and controlled designs, incorporating manipulation checks and cross-cultural sampling.

Another potential limitation concerns the skin color of the avatars. In this study, only yellow and orange skin tones were used, which may have appeared unnatural to some participants. If one color was perceived as more natural than the other, this could have biased participants' evaluations of the avatars, regardless of their facial expressions.

Finally, the experimental design focused on a binary comparison between smiling and neutral expressions. Although this approach allowed for precise control, future research should explore a wider range of emotional expressions

(e.g., genuine vs. fake smiles) using facial action coding systems to enhance the realism and granularity of the stimuli.

6.2 Future directions

For future research, it is important to investigate whether the observed under-evaluation of smiles in high-risk contexts applies exclusively to perceived friendliness, or if it extends to other social effects—such as emotional contagion or behavioral responses like smile reciprocity. Additionally, future studies should explore varying degrees and types of risk, particularly those embedded in real-life scenarios, to assess the generalizability of these findings to more ecologically valid contexts.

Furthermore, this study focused exclusively on a single non-verbal cue—smiling. To build a more comprehensive understanding of human–agent interaction, future research should consider a broader range of facial expressions and emotional cues. Investigating how different expressions (e.g., fear, sadness, neutrality, or surprise) are perceived and evaluated across varying risk levels could help clarify which cues are most context-sensitive and how they influence the emotional and cognitive responses of users.

Acknowledgment

Words cannot fully express my gratitude to my supervisor, Professor Takeuchi, for his invaluable advice and guidance. Without his insight and support, I would not have progressed as far as I have in my research and academic journey. His deep commitment to educational excellence and meticulous attention to detail have profoundly shaped the development of this dissertation.

I would also like to extend my sincere appreciation to the members of the Takeuchi Laboratory. Their support helped me navigate challenging moments, brainstorm experimental designs, and refine my ideas through insightful feedback. Above all, their moral support was essential throughout this process. The discussions and camaraderie I experienced with my peers at the university have enriched this research experience, making the journey both intellectually rewarding and personally fulfilling.

I am especially grateful to my family, and in particular my parents, whose unwavering belief in me sustained my motivation and perseverance. Finally, I thank my friends for encouraging when I needed it most, and for offering light-hearted distractions that helped me maintain balance along the way.

Reference

[1] Paul Ruvolo, Daniel Messinger, and Javier Movellan. Infants time their smiles to make their moms smile. *PLoS one*, 10(9):e0136492, 2015.

[2] Marc H Bornstein and Martha E Arterberry. Recognition, discrimination and categorization of smiling by 5-month-old infants. *Developmental Science*, 6(5):585–599, 2003.

[3] Paul Ekman. *Telling lies: Clues to deceit in the marketplace, politics, and marriage (revised edition)*. WW Norton & Company, 2009.

[4] Paul Ekman and Wallace V Friesen. Facial action coding system. *Environmental Psychology & Nonverbal Behavior*, 1978.

[5] Kuba Krys, Christin Melanie Vauclair, Colin A Capaldi, Vivian Miu-Chi Lun, Michael Harris Bond, Alejandra Domínguez-Espinosa, Claudio Torres, Ottmar V Lipp, L Sam S Manickam, Cai Xing, et al. Be careful where you smile: Culture shapes judgments of intelligence and honesty of smiling individuals. *Journal of nonverbal behavior*, 40:101–116, 2016.

[6] Theodore M Newcomb, Ralph H Turner, and Philip E Converse. *Social psychology: The study of human interaction*. Psychology Press, 2015.

[7] Thomas E Malloy and Linda Albright. Interpersonal perception in a social context. *Journal of Personality and Social Psychology*, 58(3):419, 1990.

[8] Zeynep Cemalcilar, Lemi Baruh, Murat Kezer, Roza Gizem Kamiloglu, and Bihter Nigdeli. Role of personality traits in first impressions: An investigation of actual and perceived personality similarity effects on interpersonal attraction across communication modalities. *Journal of Research in Personality*, 76:139–149, 2018.

[9] Paul Ekman and Wallace V Friesen. Felt, false, and miserable smiles. *Journal of nonverbal behavior*, 6(4):238–252, 1982.

[10] Verlin B Hinsz and Judith A Tomhave. Smile and (half) the world smiles with you, frown and you frown alone. *Personality and Social Psychology Bulletin*, 17(5):586–592, 1991.

[11] Barbara Wild, Michael Erb, Michael Eyb, Mathias Bartels, and Wolfgang Grodd. Why are smiles contagious? an fmri study of the interaction between perception of facial affect and facial movements. *Psychiatry Research: Neuroimaging*, 123(1):17–36, 2003.

[12] Patricia B Barger and Alicia A Grandey. Service with a smile and encounter satisfaction: Emotional contagion and appraisal mechanisms. *Academy of management journal*, 49(6):1229–1238, 2006.

[13] Evania L Fasya, Esther van den Bos, Dirk KJ Heylen, and Mariska E Kret. Smile mimicry smoothens human-virtual human interactions. *International Journal of Human-Computer Studies*, 183:103182, 2024.

[14] Samuele Centorrino, Elodie Djemai, Astrid Hopfensitz, Manfred Milinski, and Paul Seabright. Honest signaling in trust interactions: Smiles rated as genuine induce trust and signal higher earning opportunities. *Evolution and Human Behavior*, 36(1):8–16, 2015.

[15] Liuba Y Belkin and Naomi B Rothman. Do i trust you? depends on what you feel: Interpersonal effects of emotions on initial trust at zero-acquaintance. *Negotiation and Conflict Management Research*, 10(1):3–27, 2017.

[16] Lucy Johnston, Lynden Miles, and C Neil Macrae. Why are you smiling at me? social functions of enjoyment and non-enjoyment smiles. *British Journal of Social Psychology*, 49(1):107–127, 2010.

[17] Paul Ekman, Wallace V Friesen, and Maureen O'sullivan. Smiles when lying. *Journal of personality and social psychology*, 54(3):414, 1988.

[18] Byron Reeves and Clifford Nass. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK*, 10(10):19–36, 1996.

[19] Nicole Krämer, Stefan Kopp, Christian Becker-Asano, and Nicole Sommer. Smile and the world will smile with you—the effects of a virtual agent ‘ s smile on users ’ evaluation and behavior. *International Journal of Human-Computer Studies*, 71(3):335–349, 2013.

[20] Soo Youn Oh, Jeremy Bailenson, Nicole Krämer, and Benjamin Li. Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PloS one*, 11(9):e0161794, 2016.

[21] Ella Glikson, Arik Cheshin, and Gerben A van Kleef. The dark side of a smiley: Effects of smiling emoticons on virtual first impressions. *Social Psychological and Personality Science*, 9(5):614–625, 2018.

[22] Masahiro Mori, Karl F MacDorman, and Norri Kageki. The uncanny valley [from the field]. *IEEE Robotics & automation magazine*, 19(2):98–100, 2012.

[23] Tom Geller. Overcoming the uncanny valley. *IEEE computer graphics and applications*, 28(4):11–17, 2008.

[24] Angela Tinwell, Mark Grimshaw, Debbie Abdel Nabi, and Andrew Williams. Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in Human behavior*, 27(2):741–749, 2011.

[25] Harry T Reis, Ilona McDougal Wilson, Carla Monestere, Stuart Bernstein, Kelly Clark, Edward Seidl, Michelle Franco, Ezia Gioioso, Lori Freeman, and Kimberly Radoane. What is smiling is beautiful and good. *European Journal of Social Psychology*, 20(3):259–267, 1990.

[26] Robert E Kraut and Robert E Johnston. Social and emotional messages of smiling: an ethological approach. *Journal of personality and social psychology*, 37(9):1539, 1979.

[27] Lee Ellis. Gender differences in smiling: An evolutionary neuroandrogenic theory. *Physiology & Behavior*, 88(4-5):303–308, 2006.

[28] Thorsten Hennig-Thurau, Markus Groth, Michael Paul, and Dwayne D Gremler. Are all smiles created equal? how emotional contagion and emo-

tional labor affect service relationships. *Journal of marketing*, 70(3):58–73, 2006.

[29] Daniel S Messinger, Alan Fogel, and K Laurie Dickson. All smiles are positive, but some smiles are more positive than others. *Developmental psychology*, 37(5):642, 2001.

[30] Mehwish Khan, Syed Murtaza Raza Kazmi, Farhan Raza Khan, and Imran Samejo. Analysis of different characteristics of smile. *BDJ open*, 6(1):6, 2020.

[31] Zara Ambadar, Jeffrey F Cohn, and Lawrence Ian Reed. All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of nonverbal behavior*, 33:17–34, 2009.

[32] Jared Martin, Magdalena Rychlowska, Adrienne Wood, and Paula Niedenthal. Smiles as multipurpose social signals. *Trends in cognitive sciences*, 21(11):864–877, 2017.

[33] Paul Ekman and Erika L Rosenberg. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)*. Oxford University Press, USA, 1997.

[34] Baruch Fischhoff. Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human perception and performance*, 1(3):288, 1975.

[35] Paul Slovic, Melissa L Finucane, Ellen Peters, and Donald G MacGregor. The affect heuristic. *European journal of operational research*, 177(3):1333–1352, 2007.

[36] Shelly Chaiken and Yaacov Trope. *Dual-process theories in social psychology*. Guilford Press, 1999.