To Tell the Truth: Virtual Agents and Morning Morality

Sharon Mozgai ✉, Gale Lucas, and Jonathan Gratch

University of Southern California,
Institute for Creative Technologies, USA
mozgai@ict.usc.edu

Abstract. This paper investigates the impact of time of day on truthfulness in human-agent interactions. Time of day has been found to have important implications for moral behavior in human-human interaction. Namely, the morning morality effect shows that people are more likely to act ethically (i.e., tell fewer lies) in the morning than in the afternoon. Based on previous work on disclosure and virtual agents, we propose that this effect will not bear out in human-agent interactions. Preliminary evaluation shows that individuals who lie when engaged in multi-issue bargaining tasks with the Conflict Resolution Agent, a semi-automated virtual human, tell more lies to human negotiation partners than virtual agent negotiation partners in the afternoon and are more likely to tell more lies in the afternoon than in the morning when they believe they are negotiating with a human. Time of day does not have a significant effect on the amount of lies told to the virtual agent during the multi-issue bargaining task.

Keywords: morning morality, virtual humans, honest responding, multi-issue bargaining

1 Introduction

If you really want the truth, should you ask for it, via virtual agent, in the morning? As the number of applications of virtual agents grows, particularly in domains that depend on honest disclosure for success (e.g., healthcare and experiential learning), it is important to consider what conditions lead people to act unethically when engaging with virtual agents and what advantages virtual agents may have in eliciting honest and ethical behaviors. In addition, interest in how people respond to intelligent virtual agents versus how they respond to people has increased considerably [6, 4]. In human-human interaction, time of day has been shown to have important implications for ethical behavior. Specifically, the *morning morality effect* suggests that time of day can lead to a systematic failure of "good people" to act morally showing that individuals engage in less ethical behavior (i.e., more lying) on tasks performed in the afternoon than on tasks performed in the morning [3]. This current work considers whether the morning morality effect impacts human-agent interactions and posits that one distinct advantage of virtual agents may be inoculation against the morning morality effect.

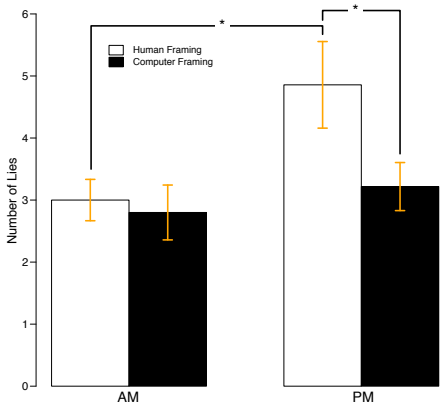

Fig. 1. Comparison between number of lies with respect to both AM and PM as well as human and computer framing conditions. Visualized are mean values with standard error bars. Brackets with * indicate significant differences at $p < 0.05$.

Initial theories regarding human-computer interaction maintained it was possible to replicate any psychological finding on how people interact with one another by replacing the word human with computer [8]. Subsequent work has shown that individuals interact differently with virtual agents, particularly in social settings, where they can exhibit lower social presence [1], arousal [8], and engagement [7]. But recent research has also identified contexts in which virtual agents outperform humans, specifically in overcoming barriers to obtaining truthful information. Lucas et al. [5] have shown that when people believe that a virtual agent was automatically-operated by a computer algorithm rather than tele-operated by a human they reported lower fear of self-disclosure, lower impression management and were rated by observers as being more willing to disclose truthfully.

We contribute to this work on obtaining truthful information via virtual agents by examining the impact of time of day on human-human and human-agent interaction in a multi-issue bargaining task. We predict individuals who lie will be less deceitful with their bargaining partner regardless of time of day when they believe they are interacting with an autonomous virtual agent than when they believe they are interacting with a human. Moreover, we predict that the morning morality effect can still be found when individuals who lie believe they are bargaining with another human but will be diminished when they believe they are negotiating with an autonomous virtual agent.

2 Current Work

This analysis was conducted on the Conflict Resolution Agent (CRA) dataset made available by the authors [2]. The CRA is a semi-automated virtual human that negotiates with people via natural language. The CRA dataset is comprised of ninety three participants (52 female) who were recruited from an on-line job service and who completed two multi-issue negotiations. All participants interacted with both the male and female versions of the virtual agent. Following a

wizard-of-Oz paradigm, all participants interacted with an agent that was controlled by a human wizard. Further, participants were randomly assigned to a framing condition, where they were either told that the agent was operated by a human or a computer (artificial intelligence). The wizards followed a predetermined script: the wizard acted as if the participants preferences were unknown, the wizard avoided making the first offer and the wizard avoided volunteering their own preferences unless the participants used reciprocal information exchange. This script was followed across both framing conditions and both negotiations. As an incentive, participants received tickets to a $100 lottery based on how well they performed in the negotiations. The number of tickets each participant received was based on the value of the items they won in the negotiation. The data was transcribed, time stamped, and annotated for lie counts.

3 Results

Out of the 93 participants 12 were not annotated for lies and participants that did not lie at all were excluded from the analyses. In total this leaves 64 participants for the evaluations shown below. Sessions were classified as morning (between 9 a.m. and noon) or afternoon (between noon and 6 p.m).

Correlation analyses results. We observe a significant correlation with respect to number of lies told and time of day (Pearson's r = 0.296; n = 64; p = 0.017). The effect is carried by the human framing (r = 0.366; n = 31; $p < 0.05$) rather than the computer framing (r = 0.204; n = 33; p = 0.25). In other words, the effect of morning morality is stronger in the human framing.

Morning vs. Afternoon. Similar to the correlation analyses we observe a significant difference in observed lies with respect to AM vs. PM participants. Specifically, subjects participating in the afternoon ($M = 4$, $SD = 2.685$) lie more than those participating in the morning ($M = 2.9$, $SD = 1.210$; t(62) = -2.256, p = 0.027, d = -0.471). Further, we observe that the effect is again carried by the human framing condition (AM: $M = 3$, $SD = 1.054$ vs. PM: $M = 4.857$, $SD = 3.198$; t(29) = -2.401, p = 0.024, d = 0.936). We do not observe a significant effect for the amount of lies expressed in the computer framing between morning and afternoon (AM: $M = 2.8$, $SD = 1.398$ vs. PM: $M = 3.217$, $SD = 1.858$; t(31) = -0.710, p = 0.485).

Computer vs. Human Framing. While the result is not significant, participants seem to lie more in the human framing condition ($M = 4.258$, $SD = 2.816$) than the computer framing condition ($M = 3.091$, $SD = 1.721$; t(62) = 1.986, p = 0.053, d = -0.497). When comparing the lies between the framing conditions during the morning and afternoon separately we see a significant difference for the afternoon (human framing: $M = 4.857$, $SD = 3.198$ vs. computer framing: $M = 3.217$, $SD = 1.858$; t(42) = 2.054, p = 0.048, d = 0.635) while there is no difference in the morning (human framing: $M = 3$, $SD = 1.054$ vs. computer framing: $M = 2.8$, $SD = 1.398$; t(18) = 0.361, p = 0.722). The results are summarized in Figure 1.

4 Conclusions

As previous work on disclosure and virtual agents has found, people seem to be more truthful when interacting with virtual humans than when interacting with other people. The current work extends this research to show that not only are people more truthful with virtual agents, they are more truthful despite situational variables, namely time of day, that can significantly affect ethical behavior. This is useful information regarding the utility of virtual agents in contexts where honest disclosure is essential. If people behave more or less ethically depending on time of day, employing a virtual agent for truthful disclosure, particularly in the afternoon and evening, may be preferred to overcome barriers to honesty. Additional inquiry is necessary to fully unpack what might be mediating this phenomenon. One potential theory is that humans engaged in a task with a virtual agent assume the agent has more or perfect information and therefore determine that any lie will be in vain. There might also be an effect of novelty; as people become more accustomed to interacting with virtual agents, this effect may extinguish. Longitudinal investigation is suggested to evaluate whether or not the novelty of interacting with a virtual agent can account for these behaviors.

Acknowledgements
This work was supported by the National Science Foundation under grants IIS-1211064 and BCS-1419621, and the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

1. Gajadhar, B., de Kort, Y., IJsselsteijn, W.: Shared fun is doubled fun: player enjoyment as a function of social setting. Fun and games pp. 106–117 (2008)
2. Gratch, J., DeVault, D., Lucas, G.: The benefits of virtual humans for teaching negotiation. In: International Conference on Intelligent Virtual Agents. pp. 283–294. Springer (2016)
3. Kouchaki, M., Smith, I.H.: The morning morality effect: The influence of time of day on unethical behavior. Psychological science 25(1), 95–102 (2014)
4. Krämer, N.C., von der Pütten, A., Eimler, S.: Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In: Human-computer interaction: The agency perspective, pp. 215–240. Springer (2012)
5. Lucas, G.M., Gratch, J., King, A., Morency, L.P.: Its only a computer: virtual humans increase willingness to disclose. Computers in Human Behavior 37, 94–100 (2014)
6. de Melo, C.M., Gratch, J.: Beyond believability: quantifying the differences between real and virtual humans. In: International Conference on Intelligent Virtual Agents. pp. 109–118. Springer (2015)
7. Ravaja, N.: The psychophysiology of digital gaming: The effect of a non co-located opponent. Media Psychology 12(3), 268–294 (2009)
8. Reeves, B., Nass., C.: The media equation: How people treat computers, television, and new media like real people and places. Cambridge University Press (1996)