Stop Copying Me: Evaluating nonverbal mimicry in embodied motivational agents

Isaac Wang
wangid@jmu.edu
James Madison University
Harrisonburg, Virginia, USA

Rodrigo Calvo, Heting Wang, Jaime Ruiz
rodrigo.calvo@ufl.edu,heting.wang@ufl.edu,jaime.ruiz@ufl.edu
University of Florida
Gainesville, Florida, USA

ABSTRACT

Motivational agents are virtual agents that seek to motivate users by providing feedback and guidance. Prior work has shown how certain factors of an agent, such as the type of feedback given or the agent's appearance, can influence user motivation when completing tasks. However, it is not known how nonverbal mirroring affects an agent's ability to motivate users. Specifically, would an agent that mirrors be more motivating than an agent that does not? Would an agent trained on real human behaviors be better? We conducted a within-subjects study asking 30 participants to play a "find-the-hidden-object" game while interacting with a motivational agent that would provide hints and feedback on the user's performance. We created three agents: a Control agent that did not respond to the user's movements, a simple Mimic agent that mirrored the user's movements on a delay, and a Complex agent that used a machine-learned behavior model. We asked participants to complete a questionnaire asking them to rate their levels of motivation and perceptions of the agent and its feedback. Our results showed that the Mimic agent was more motivating than the Control agent and more helpful than the Complex agent. We also found that when participants became aware of the mimicking behavior, it can feel weird or creepy; therefore, it is important to consider the detection of mimicry when designing virtual agents.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **Gestural input**; *Human computer interaction (HCI)*.

KEYWORDS

virtual agents, motivational agents, behavioral mirroring, nonverbal

ACM Reference Format:
Isaac Wang and Rodrigo Calvo, Heting Wang, Jaime Ruiz. 2023. Stop Copying Me: Evaluating nonverbal mimicry in embodied motivational agents. In *ACM International Conference on Intelligent Virtual Agents (IVA '23), September 19–22, 2023, Würzburg, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3570945.3607322

ACM ISBN 978-1-4503-9994-4/23/09.
https://doi.org/10.1145/3570945.3607322

Figure 1: Example of our study setup. The user (right) is playing a find-the-hidden-object game while interacting with a motivational agent (left) that gives them hints and feedback during the game.

1 INTRODUCTION AND RELATED WORK

Motivational agents are agents that provide feedback to help guide and motivate users. They often monitor users while they complete tasks and give specific feedback on their performance [17, 20, 24]. For example, a motivational agent by Mumm and Mutlu [20] was able to increase intrinsic motivation in users and also increase their overall task performance, pointing to the positive effects that motivational agents can have on users. A number of different factors can influence the efficacy of a motivational agent, such as the agent's dialogue and feedback [8, 16, 20, 26], but also nonverbal factors such as appearance [4, 12, 15] and behavior [18, 21]. However, it is not known how nonverbal mirroring (i.e., mimicking a user's motions) can be used in motivational agents. Mirroring has the potential to increase feelings of likeability and persuasion [1, 10], which could also influence motivation [32].

Thus, the goal of our work is to investigate how mirroring can be used in motivational agents and how it affects the way users perceive an agent and its feedback. Specifically, would an agent that mirrors a user's motions be more motivating than an agent that does not? Furthermore, would natural data-driven behaviors be more motivating since it models real human motion? In this paper, we present a study asking participants to complete a find-the-hidden-object game while interacting with an agent that would give hints and provide feedback on their performance.

We created three agents that each exhibited different nonverbal behaviors: a *Control* agent that looked around randomly and did not respond to users' movements, a *Mimic* agent that displayed simple gaze and expression mimicry on a delay, and a *Complex* agent that used a data-driven model of real human behavior to respond to the user's movements and expressions. We evaluated each agent to explore how mirroring can affect user perceptions and motivation, providing researchers with new insights on how to design motivational agents.

2 METHOD

We conducted a within-subjects study to understand how agent behavior affects users' perceptions of the agent and users' levels of motivation. We used a game as our experimental task, similar to prior motivational studies [19, 22]. The goal of each level in the game was to find an object hidden in an image while interacting with a virtual agent [29]. We had a total of 18 different levels, which were randomized for each participant. Participants were seated in front of a touchscreen computer and interacted with the game and virtual agent. The agent's speech responses were controlled via a wizard-of-oz interface.

During each level, a virtual agent positioned next to the image would accompany the user, serving as an assistant. While searching for the object, users could talk to the agent and ask for hints. Our agent supported four hints for each object, the object's color, the context of what is around the object, the shape of the object, and a general location of where the object is in the image. After each level, the agent would encourage the user regardless of whether they found the object.

2.1 Agent Behaviors

The agents and game were developed in Unity. OpenFace [3] was used for tracking users' motions and FACSvatar [27] was used to drive the agents' motions. We developed three different agents (*Control*, *Mimic*, and *Complex*), each exhibiting different nonverbal behaviors. Each agent had the same appearance and speech; the only difference was their nonverbal behavior.

The *Control* agent would look randomly around and vary its gaze every few seconds, simulating the act of searching for the object alongside the user. For facial expressions, the agent would smile as it delivered feedback; otherwise, the agent only moved its head around during the game and when talking to the user. This agent's movement did not correspond to the user's movement.

For the *Mimic* agent, we used a head-gaze mirroring technique originally proposed by Bailenson et al. [1]. The agent would track the user's head movements and mirror them on a delay, essentially mimicking the user's gaze. We implemented a four-second delay as prior work showed delay between 3-4s is appropriate [14, 23, 25]. The agent also tracked the user's smile and mimicked it alongside the head movements with the same four-second delay.

For the *Complex* agent, we utilized the IL-LSTM model by Dermouche and Pelachaud [11]. The machine learning model was trained on the NoXi [9] dataset, consisting of facial tracking data from human-human interactions between an expert and a novice. The model would take in the user's head gaze and smile expression and predict the agent's corresponding head gaze and smile expression. We chose this model as it fit the interaction between user and agent in our game, with the user playing the role of a novice (who is looking for objects) and the agent playing the role of an expert (who knows where the objects are and is trying to guide the novice). This model has the potential to be a better fit for a motivational agent's behavior than the simple delayed mimicry, as a model trained on real data should better reflect natural behavior when responding to a user's movements and expressions.

2.2 Study

We recruited 33 participants (17 female and 16 male) from computer science classes at a local university. Participants were between the ages of 20 to 35 (mean=23). Out of the 33 participants, 31 had prior experience with voice assistants (e.g., Siri, Alexa). Participants received extra credit in a course as compensation. Our study was approved by our Institutional Review Board.

At the start of the study, participants were seated in front of the touchscreen and the researcher would then describe the game and have them complete a practice level. We did not describe the differences between the agents nor how they would/would not respond to participants' motions. Participants would then complete three trials of the study. For each trial, participants would interact with one of the three agents and complete six levels of the game. They would then fill out a questionnaire consisting of the Situational Motivation Scale (SIMS) [13] to measure their levels of motivation, the Agent Rating Questionnaire (ARQ) [28] to rate their perceptions of the agent, and the Positive and Negative Affect Schedule (PANAS) [30] to measure their levels of affect. This process was repeated for the next two agents. At the end of the study, we conducted a short interview asking participants which agent they liked most/least, how their levels of motivation changed throughout the study, and if they noticed anything specific about the agents' behaviors. The order of the agents was counterbalanced using a full counterbalanced design.

3 RESULTS AND DISCUSSION

In this section, we discuss our results from the questionnaires and semi-structured interviews. Out of the 33 participants, we omitted data from 3 participants (P1, P2, P29) due to system malfunctions, leaving data from 30 participants for analysis. As part of our analysis, we first ran a Shapiro-Wilks test for normality. If the distribution was not normal, we applied an Aligned Rank Transform (ART) [31] to the data before conducting a repeated measures analysis of variance (ANOVA) with Behavior and Trial as the independent (within-subject) variables.

3.1 User Motivation

To understand how agent behavior affects user motivation, we analyzed the results from the SIMS questionnaire. This included totaling the responses for the 16 questions against the four aspects of motivation measured by the SIMS. An ART–ANOVA revealed no significant effect of Behavior on the categories of *Intrinsic Motivation* ($F_{2,52}$ = 1.02, n.s.), *Identified Regulation* ($F_{2,52}$ = .02, n.s.), or *External Regulation* ($F_{2,52}$ = .26, n.s.). However, ART-ANOVA did reveal a significant main effect of Behavior on *Amotivation* ($F_{2,52}$ = 3.95, p<.05). Post-hoc T-tests with Tukey HSD correction showed that participants had significantly higher feelings of *Amotivation* with the *Control* agent (M=9.33, SD=5.23) compared to the *Mimic* agent (M=8.63, SD=4.80), showing that ratings were on average 0.7 points higher for the *Control* agent. This indicated participants were less motivated with the *Control* agent when compared to the *Mimic* agent, with the *Complex* agent in between (but not significantly different from the others). Based on our qualitative results, the reason might be the *Control* agent had behaviors unrelated to the user's movements and lack of eye contact/attention and mirroring behaviors. This may have been demotivating to participants. Ratings for the *Complex* agent (M=8.93, SD=5.79) were not significantly different from the other two. We did not observe any main nor interaction effects for TRIAL.

Overall, the ability of the Mimic agent to maintain participants' motivation levels shows promise for mirroring behaviors in motivational agents. We recommend balancing the motivational agent's nonverbal behavior: the agent should be responsive to the user's behavior without acting as if it is completely mimicking the user. The Mimic agent achieves some balance by mirroring user motions without training. However, it still needs improvement in communicating sufficiently with the user and resuming tasks. This can be done by establishing a shared gaze while displaying general gaze and expression mirroring. Specifically, future research could design an agent that recognizes and responds the user's intent and switch between different behaviors.

3.2 Agent Perceptions and Detection of Mimicry

We analyzed the ARQ results to see if user perceptions of an agent differed based on the agent's nonverbal behavior. An ART-ANOVA revealed no significant effect of agent BEHAVIOR on the overall cumulative scores ($F_{2,52}$ = .06, n.s.); however, when analyzing the individual scales, ART-ANOVA did reveal a significant main effect of BEHAVIOR on ratings of *Helpfulness* (F2,52=11.97, p<.001). Post-hoc T-tests with Tukey HSD correction showed that participants rated the *Complex* agent as having lower *Helpfulness* (M=4.27, SD=.87) compared to both the *Control* (M=4.47, SD=.63) and the *Mimic* (M=4.57, SD=.73) agents. On average, the *Complex* agent was rated 0.25 points lower than the other two agents. However, the only difference between the agents was their nonverbal behavior. Thus, it appears that the Complex agent's nonverbal behaviors influenced participants' perceptions of the agent's helpfulness.

In our qualitative analysis, we saw that participants interpreted the Complex agent as mimicking them, even more so than the Mimic agent. The Complex agent responds to the user's motion more immediately than the Mimic agent, which could explain why participants interpreted it as mimicry more often. Previous work demonstrated that the detection of mimicry was correlated with discomfort, a sense of distrust, and lower ratings in terms of trustworthiness and warmth [2, 6, 7]. We saw similar comments from our participants, with some describing the mimicry as awkward, weird, and even creepy. We believe that, in this case, the detection of mimicry was influential enough to change participants' perceptions. Thus, it is important to consider the detection of mimicry when creating models of nonverbal behavior.

We also found a significant main effect of TRIAL on *Helpfulness* ($F_{2,52}$ = 7.96, p<.001). Post-hoc tests showed that *Helpfulness* was significantly higher for the first trial (M=4.53, SD=.73) compared against both the second (M=4.40, SD=.56) and third (M=4.37, SD=.93) trials. We found similar themes in our qualitative data; participants mentioned how the three agents felt the same and that the hints and feedback were repetitive across all of them. As participants progressed through the study, they may have felt more capable at the game, leading them to perceive subsequent hints as less helpful. This change in helpfulness could also be attributed to the phenomenon of impressions of an agent evolving over time [5]. Ratings for the second and third trials were not significantly different from each other, and we did not observe any interaction effects between TRIAL and BEHAVIOR.

3.3 Affective Reactions

We examined participants' affective responses to agent feedback by analyzing responses to the PANAS questionnaire. We computed mean *Positive Affect* and *Negative Affect* score based on the 20 terms in the questionnaire. An ART-ANOVA revealed no significant effect of agent BEHAVIOR on either *Positive Affect* ($F_{2,52}$ = .49, n.s.) or *Negative Affect* ($F_{2,52}$ = 1.65, n.s.).

However, we did find a significant main effect of TRIAL on *Positive Affect* ($F_{2,52}$ = 3.92, p<.05). Post-hoc T-tests with Tukey HSD correction showed that *Positive Affect* was lowest for the third trial ($M = 14.07, SD = 5.94$) compared to both the first (M=15.53, SD=6.48) and second (M=14.83, SD=6.09) trials. The decrease in positive affective responses to agent feedback in the third trial was supported by participant comments highlighting the repetitive nature of the feedback. However, despite this decline, there was no noticeable decrease in motivation or other study outcomes over time. No significant difference was found between the first and second trials, and we did not observe any interaction effects; however, it may be worth considering varying the feedback provided by the agent to maintain positive user reactions throughout the study.

4 CONCLUSION AND FUTURE WORK

In this paper, we presented a study investigating how different types of responsive nonverbal behaviors influence user perceptions of a motivational and feedback-giving agent. We asked participants to interact with three different agents, each with different nonverbal behavior: a Control agent that did not respond to the user's movements, a Mimic agent that mimicked the gaze and expressions of the user, and a Complex agent that used a data-driven model of real human behavior to respond to the user's movements and expressions. We showed that the Mimic agent was able to avoid demotivating participants while remaining helpful. In addition, we showed how the detection of mimicry needs to be considered when designing agent behaviors as it may negatively impact user perceptions. In our study, we used a hidden-object game with a user completing tasks while an agent assisted them and gave feedback on the side. Our results may not be generalizable to all domains; however, future work should focus on similar tasks, such as in interactive learning systems [16, 26], in which a user would complete practice problems with an agent that coaches them. Overall, our findings help inform the design of nonverbal behaviors for motivational and feedback-giving agents.

ACKNOWLEDGMENTS

This work is partially supported by the National Science Foundation Grant Award #IIS-1750840. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect these agencies' views.

REFERENCES

[1] Jeremy N. Bailenson and Nick Yee. 2005. Digital Chameleons: Automatic Assimilation of Nonverbal Gestures in Immersive Virtual Environments. *Psychological Science* 16, 10 (Oct 2005), 814–819. https://doi.org/10.1111/j.1467-9280.2005.01619.x

[2] Jeremy N. Bailenson, Nick Yee, Kayur Patel, and Andrew C. Beall. 2008. Detecting digital chameleons. *Computers in Human Behavior* 24, 1 (Jan 2008), 66–87. https://doi.org/10.1016/j.chb.2007.01.015

[3] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. OpenFace 2.0: Facial Behavior Analysis Toolkit. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. IEEE, 59–66. https://doi.org/10.1109/FG.2018.00019

[4] Amy L. Baylor. 2011. The design of motivational agents and avatars. *Educational Technology Research and Development* 59, 2 (Apr 2011), 291–300. https://doi.org/10.1007/s11423-011-9196-3

[5] 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 MarilynEditors Walker (Eds.). Springer Berlin Heidelberg, 126–138.

[6] Quinten S. Bernhold and Howard Giles. 2020. Vocal Accommodation and Mimicry. *Journal of Nonverbal Behavior* 44, 1 (Mar 2020), 41–62. https://doi.org/10.1007/s10919-019-00317-y

[7] David B. Buller and Judee K. Burgoon. 1996. Interpersonal Deception Theory. *Communication Theory* 6, 3 (1996), 203–242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x

[8] Christian Burgers, Allison Eden, Mélisande D. van Engelenburg, and Sander Buningh. 2015. How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior* 48 (Jul 2015), 94–103. https://doi.org/10.1016/j.chb.2015.01.038

[9] Angelo Cafaro, Johannes Wagner, Tobias Baur, Soumia Dermouche, Mercedes Torres Torres, Catherine Pelachaud, Elisabeth André, and Michel Valstar. 2017. The NoXi database: multimodal recordings of mediated novice-expert interactions. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI '17)*. Association for Computing Machinery, 350–359. https://doi.org/10.1145/3136755.3136780

[10] Tanya L. Chartrand and John A. Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology* 76, 6 (1999), 893. https://doi.org/10.1037/0022-3514.76.6.893

[11] Soumia Dermouche and Catherine Pelachaud. 2019. Generative Model of Agent's Behaviors in Human-Agent Interaction. In *2019 International Conference on Multimodal Interaction (ICMI '19)*. Association for Computing Machinery, 375–384. https://doi.org/10.1145/3340555.3353758

[12] Dan Feng, David C. Jeong, Nicole C. Krämer, Lynn C. Miller, and Stacy Marsella. 2017. "Is It Just Me?": Evaluating Attribution of Negative Feedback as a Function of Virtual Instructor's Gender and Proxemics. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS '17)*. International Foundation for Autonomous Agents and Multiagent Systems, 810–818.

[13] Frédéric Guay, Robert J. Vallerand, and Céline Blanchard. 2000. On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion* 24, 3 (Sep 2000), 175–213. https://doi.org/10.1023/A:1005614228250

[14] Nicole Krämer, Stefan Kopp, Christian Becker-Asano, and Nicole Sommer. 2013. 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 (Mar 2013), 335–349. https://doi.org/10.1016/j.ijhcs.2012.09.006

[15] Nicole C. Krämer, Bilge Karacora, Gale Lucas, Morteza Dehghani, Gina Rüther, and Jonathan Gratch. 2016. Closing the gender gap in STEM with friendly male instructors? On the effects of rapport behavior and gender of a virtual agent in an instructional interaction. *Computers & Education* 99 (Aug 2016), 1–13. https://doi.org/10.1016/j.compedu.2016.04.002

[16] Lijia Lin, Robert K. Atkinson, Robert M. Christopherson, Stacey S. Joseph, and Caroline J. Harrison. 2013. Animated agents and learning: Does the type of verbal feedback they provide matter? *Computers & Education* 67 (Sep 2013), 239–249. https://doi.org/10.1016/j.compedu.2013.04.017

[17] Kristen E. Link, Roger J. Kreuz, Arthur C. Graesser, and The Tutoring Research Group. 2001. Factors that Influence the Perception of Feedback Delivered by a Pedagogical Agent. *International Journal of Speech Technology* 4, 2 (Jun 2001), 145–153. https://doi.org/10.1023/A:1017383528041

[18] Richard E. Mayer and C. Scott DaPra. 2012. An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied* 18, 3 (2012), 239–252. https://doi.org/10.1037/a0028616

[19] Naoki Miura, Hiroki C Tanabe, Akihiro T Sasaki, Tokiko Harada, and Norihiro Sadato. 2017. Neural evidence for the intrinsic value of action as motivation for behavior. *Neuroscience* 352 (2017), 190–203.

[20] Jonathan Mumm and Bilge Mutlu. 2011. Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback. *Computers in Human Behavior* 27, 5 (Sep 2011), 1643–1650. https://doi.org/10.1016/j.chb.2011.02.002

[21] Ali Oker, Florian Pecune, and Christelle Declercq. 2020. Virtual tutor and pupil interaction: A study of empathic feedback as extrinsic motivation for learning. *Education and Information Technologies* 25, 5 (Sep 2020), 3643–3658. https://doi.org/10.1007/s10639-020-10123-5

[22] Eleanor O'Rourke, Kyla Haimovitz, Christy Ballweber, Carol Dweck, and Zoran Popović. 2014. Brain points: A growth mindset incentive structure boosts persistence in an educational game. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 3339–3348.

[23] Stéphane Raffard, Robin N. Salesse, Catherine Bortolon, Benoit G. Bardy, José Henriques, Ludovic Marin, Didier Stricker, and Delphine Capdevielle. 2018. Using mimicry of body movements by a virtual agent to increase synchronization behavior and rapport in individuals with schizophrenia. *Scientific Reports* 8, 1 (Nov 2018), 17356. https://doi.org/10.1038/s41598-018-35813-6

[24] Jong-Eun Roselyn Lee, Clifford Nass, Scott Brenner Brave, Yasunori Morishima, Hiroshi Nakajima, and Ryota Yamada. 2007. The Case for Caring Colearners: The Effects of a Computer-Mediated Colearner Agent on Trust and Learning. *Journal of Communication* 57, 2 (Jun 2007), 183–204. https://doi.org/10.1111/j.1460-2466.2007.00339.x

[25] Daniel Roth, David Mal, Christian Felix Purps, Peter Kullmann, and Marc Erich Latoschik. 2018. Injecting Nonverbal Mimicry with Hybrid Avatar-Agent Technologies: A Naive Approach. In *Proceedings of the Symposium on Spatial User Interaction (SUI '18)*. Association for Computing Machinery, 69–73. https://doi.org/10.1145/3267782.3267791

[26] Su-Mae Tan, Tze Wei Liew, and Chin Lay Gan. 2020. Motivational virtual agent in e-learning: the roles of regulatory focus and message framing. *Information and Learning Sciences* 121, 1/2 (Jan 2020), 37–51. https://doi.org/10.1108/ILS-09-2019-0088

[27] Stef van der Struijk, Hung-Hsuan Huang, Maryam Sadat Mirzaei, and Toyoaki Nishida. 2018. FACSvatar: An Open Source Modular Framework for Real-Time FACS based Facial Animation. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents (IVA '18)*. Association for Computing Machinery, 159–164. https://doi.org/10.1145/3267851.3267918

[28] Isaac Wang, Lea Buchweitz, Jesse Smith, Lara-Sophie Bornholdt, Jonas Grund, Jaime Ruiz, and Oliver Korn. 2020. Wow, You Are Terrible at This! An Intercultural Study on Virtual Agents Giving Mixed Feedback. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents (IVA '20)*. Association for Computing Machinery, 1–8. https://doi.org/10.1145/3383652.3423887

[29] Isaac Wang, Jesse Smith, and Jaime Ruiz. 2019. Exploring Virtual Agents for Augmented Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, 1–12. https://doi.org/10.1145/3290605.3300511

[30] David Watson, Lee Anna Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology* 54, 6 (1988), 1063.

[31] 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*. ACM, 143–146.

[32] Shwu Ming Wu. 2012. Relationships among Perceived Likeability of Principal, School Identity, and Teacher Motivation. *Comprehensive Psychology* 1 (Jan 2012), 01.11.21.CP.1.6. https://doi.org/10.2466/01.11.21.CP.1.6

Received April 27, 2023; accepted June 22, 2023