Applied Human Factors and Ergonomics (AHFE2025), Vol. 199, 2025, 996–1005
https://doi.org/10.54941/ahfe1006912

Designing Persuasive Interactions With Pet-Type Virtual Agents: Effects of Emotion and Context in Mixed Reality

Kaoru Sumi and Rio Harada

Future University Hakodate, Hakodate, Hokkaido, Japan

ABSTRACT

Persuasive technology research has highlighted how emotionally expressive agents can encourage behavioral change in everyday life. Pet-type virtual agents, with their familiar and socially accepted forms, provide a promising model for persuasion. Our previous work showed that a four-legged pet agent expressing sadness effectively promoted compliance, whereas happiness and anger were often misinterpreted. This study extends that work by examining how emotional expressions and behavioral cues interact in mixed reality (MR) environments. A virtual pet dog was developed that combined four emotions (sadness, happiness, anger, neutral) with three behavioral actions: attention calling (approaching, barking, eye contact), guiding (moving toward a target location), and pointing (alternating gaze between an object and its destination). These combinations were tested across daily contexts, including pet-related tasks (feeding, tidying toys), non-pet-related tasks (reading, book tidying, waste disposal), limiting behaviors (reducing smartphone use), and emergency warnings (moving outside a room). Two experimental settings were compared: immersive MR interaction using HoloLens 2 and video-based presentation. Results showed that sadness with pointing was perceived as supportive, while anger with guiding often caused discomfort. Video conditions achieved higher success in simple tasks, whereas MR highlighted users' expectations for interactivity and responsiveness. These findings provide design guidelines for persuasive agents in MR, with implications for habit improvement in education, healthcare, and eldercare.

Keywords: Human-centered design, Persuasive technology, Virtual agents, Mixed reality (MR), Emotion and behavior design

INTRODUCTION

Persuasive technology aims to design interactive systems that can influence users' attitudes and behaviors to support positive change in everyday life. Since the foundational work of Fogg (2002), applications have been widely explored in domains such as health promotion, environmental sustainability, and education, often through message framing, gamification, or social feedback. More recently, research has begun to investigate how embodied agents with emotional and social capabilities can provide more natural and engaging persuasive interactions (Bickmore & Picard, 2005).

Among the various forms of agents, pet-type virtual agents are particularly promising because of their familiarity and their ability to evoke empathy and trust. People intuitively interpret the intentions of animals through simple cues such as gaze, posture, or vocalization, an ability well documented in canine communication studies (Bradshaw & Rooney, 2016; Miklósi et al., 1998; Sopроni et al., 2001). These behaviors can be repurposed as persuasive signals in human–computer interaction. Prior work demonstrated that a four-legged dog-like agent expressing sadness could effectively promote compliance, while happiness and anger were often misunderstood. These findings suggest that emotional expression plays a crucial role in shaping the persuasiveness of virtual agents.

However, emotions alone may not be sufficient for effective persuasion. In everyday life, behavioral cues such as calling for attention, guiding toward a target, or pointing to an object are equally important in communicating intentions. Furthermore, the rapid development of mixed reality (MR) technologies has created new opportunities for persuasive interaction, where virtual agents share the same physical space with users. Unlike video-based presentation, MR environments allow users to experience agents as social partners, which raises new design challenges and expectations for interactivity and responsiveness.

This study addresses these gaps by systematically examining the persuasive effects of combining emotional expressions (sadness, happiness, anger, neutral) with behavioral cues (attention calling, guiding, pointing) in different daily contexts. A dog-like virtual agent was implemented in MR and compared with a video-based condition across tasks such as feeding, tidying, reading, limiting smartphone use, and emergency evacuation. By analyzing both behavioral outcomes and subjective impressions, this work provides design guidelines for human-centered persuasive agents that can support habit improvement in education, healthcare, and eldercare.

RELATED WORK

Research on persuasive technology has explored various strategies for influencing user behavior. Fogg (2002) introduced the concept of persuasive technology, emphasizing how computers and digital systems can change what people think and do. Later frameworks such as Oinas-Kukkonen & Harjumaa (2018) refined this into models that highlight system features and processes for designing effective persuasive systems. Applications have ranged from sustainability (Midden et al., 2008) to healthcare and education.

With the rise of embodied and interactive systems, emotional and social capabilities of agents have been investigated to support long-term engagement. For example, Bickmore & Picard (2005) explored how relational agents can maintain ongoing interactions, while Breazeal (2003) examined the role of emotion in sociable humanoid robots. These studies suggest that affective and communicative features enhance persuasiveness and trust.

In particular, pet-type agents have gained attention for their familiarity and intuitive communicative cues. Work in ethology and comparative psychology has shown that dogs understand and respond to human signals such as gaze and pointing (Miklósi et al., 1998; Soproni et al., 2001). Bradshaw & Rooney

(2016) further discussed the social and communication behavior of domestic dogs, highlighting their natural ability to elicit empathy and compliance. This knowledge has inspired robotic implementations, such as Sony's AIBO (Fujita, 2001) and studies of robotic pets in care contexts (Wada & Shibata, 2007).

Gesture and multimodal communication also play an important role in embodied interaction. Cassell et al. (2000) provided a foundational framework for embodied conversational agents, while Cassell & Stone (1999), McNeill (1992), and Kendon (2004) established how gestures function as meaningful communicative acts. Building on these insights, Niewiadomski & Pelachaud (2007) proposed models for managing facial expressions in embodied agents, demonstrating the importance of synchronizing emotional and behavioral cues.

Recent advances in interaction modalities further expand persuasive potential. Mixed Reality (MR) enables agents to inhabit the same perceptual space as users, enhancing social presence. Research on gaze and proxemics has shown that spatial factors strongly influence user acceptance and engagement (Mutlu et al., 2009; Koay et al., 2014). Likewise, Lystbæk et al. (2022) demonstrated the power of combining gaze with mid-air pointing in AR menus, illustrating how multimodal cues can shape persuasive interaction.

Taken together, prior work highlights three key insights: (1) persuasive systems are strengthened when grounded in relational and emotional interaction, (2) pet-type metaphors leverage intuitive human–animal communication, and (3) MR environments offer new possibilities for socially present and persuasive agents. These insights directly inform the design and hypotheses of the present study.

DESIGN OF THE PERSUASIVE PET-TYPE AGENT

The persuasive pet-type virtual agent was designed as a four-legged dog-like character that integrates emotional expressions with behavioral actions to influence user behavior. The goal was to achieve familiarity, social acceptability, and clarity of intention in a human-centered design framework.

Four emotional states were implemented: sadness, happiness, anger, and neutral. Sadness was represented by lowered posture, reduced tail movement, and slower gaze shifts. Happiness included upright posture, energetic tail wagging, and frequent eye contact. Anger was expressed through tense posture, forward-leaning stance, and short barking, while the neutral state maintained a calm and steady stance. These emotions were selected because prior studies have shown their differential effects on compliance and user perception in persuasive contexts (Breazeal, 2003; Bickmore & Picard, 2005; Okada, 2023).

The four emotional states, Neutral, Sadness, Anger, and Joy, are represented through variations in head tilt, ear orientation, tail position, and vocalization style.

To complement these emotional cues, three types of behavioral actions were designed: attention calling, guiding, and pointing. Attention calling involved the agent approaching the user, producing soft barks, and sustaining eye contact. Guiding consisted of moving toward a target location and pausing to indicate direction. Pointing was realized through alternating gaze and head orientation between the object and the target, reflecting natural communicative strategies observed in dogs (Miklósi et al., 1998; Soproni et al., 2001).

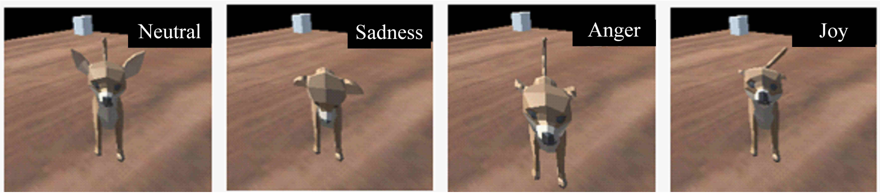

Figure 1: Emotional expressions of the dog-like agent.

Table 1: Mapping of behavioral goals to emotion–motion combinations.

Target Behavior	Motion Type	Emotion	Example Interpretation
Reading	Attention Calling	Joy	Dog playfully draws user toward a book
Trash Disposal	Pointing	Sadness	Dog looks sad while indicating trash bin
Smartphone Suppression	Guiding	Anger	Dog firmly nudges gaze away from phone
Toy Cleanup	Guiding	Sadness	Dog gently leads user to scattered toys
Book Tidying	Pointing	Confusion	Dog appears puzzled near messy bookshelf
Feeding the Dog	Attention Calling	Joy	Dog excitedly invites interaction
Emergency Relocation	Guiding	Anger	Dog urgently moves toward doorway

The system integrated these emotions and actions into twelve distinct conditions (four emotions × three actions). For example, the "sadness + pointing" combination displayed a lowered posture with alternating gaze, while the "happiness + guiding" combination showed energetic tail movements accompanied by movement toward the target. These design variations allowed systematic analysis of how different combinations influenced persuasiveness.

The virtual agent was implemented in two conditions. In the mixed reality condition, Microsoft HoloLens 2 was used to project the agent into the physical environment, enabling users to experience co-presence with the agent. In the video condition, pre-recorded animations of the agent were displayed on a standard screen to provide a baseline comparison.

The design choices were guided by three principles. First, familiarity: leveraging pet-like forms and behaviors builds on users' existing human–animal interaction schemas. Second, clarity of intention: nonverbal cues such as pointing and guiding communicate persuasive goals without verbal explanation. Third, human-centered persuasion: emotional expressions were employed not only to increase compliance but also to ensure social acceptability and minimize discomfort.

EXPERIMENTAL METHOD

Experiment 1 (Persuasive Actions in MR)

The first experiment was conducted to investigate how emotional expressions and behavioral actions influence persuasion when implemented in a dog-like virtual agent. Twenty university students (14 male, 6 female, aged 18–33) participated in this study. All participants had normal or corrected-to-normal vision and no prior experience with the virtual agent. They were randomly assigned to one of two groups. In the Proposed Method group (n = 10), the agent combined emotional expressions such as sadness or happiness with behavioral actions such as approaching, gazing, and repeating signals. In the Bark-Only group (n = 10), the agent only barked while looking at the target object. Participants were asked to perform tasks related to pet care, such as feeding and tidying toys, when prompted by the agent. The agent appeared in the participants' physical space through a Microsoft HoloLens 2 device. Behavioral data such as task compliance were recorded, and after each trial, participants completed questionnaires evaluating perceived persuasiveness and clarity of the agent's intention.

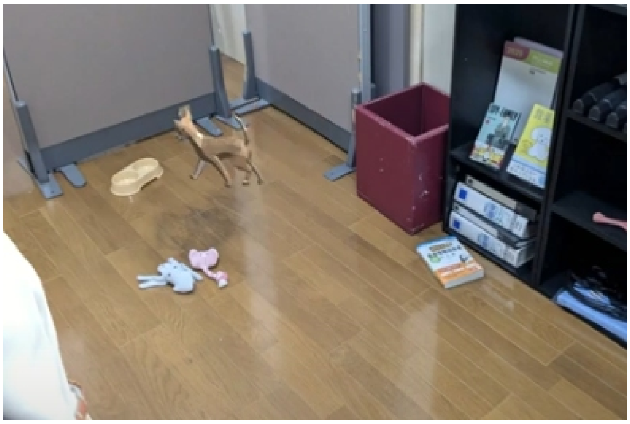

Figure 2: Experimental setup for MR (left) and video-based (right) conditions.

The MR condition used HoloLens 2 for spatially anchored interaction, while the video condition presented recorded agent behavior on a monitor from a third-person perspective.

Experiment 2 (Comparison of MR and Video)

Building on the findings of Experiment 1, a second experiment was conducted to examine the effects of modality by comparing Mixed Reality (MR) and video-based presentations. Forty university students participated in this study. Each participant experienced all twelve conditions, created by combining four emotional expressions (sadness, happiness, anger, neutral) with three behavioral actions (attention calling, guiding, pointing). The conditions were presented in randomized order, and the sequence was counterbalanced across participants to reduce order effects.

In the MR condition, the agent was projected into the participants' real environment using HoloLens 2, allowing them to experience co-presence with the agent. In the video condition, pre-recorded animations of the same agent were shown on a standard computer display. Each participant completed seven daily life scenarios, including pet-related activities (feeding, tidying toys), non-pet-related activities (reading, tidying books, disposing of waste), limiting behaviors (discouraging smartphone use), and emergency warnings (prompting the participant to move outside a room).

During each trial, the agent attempted to persuade the participant to carry out the designated task. Behavioral responses such as compliance and response time were recorded. After each condition, participants rated the persuasiveness of the agent, the emotional impression (e.g., supportive, annoying, trustworthy), the clarity of intention, and the degree of social presence. In addition, qualitative feedback was collected through open-ended questions.

For both experiments, data analysis included compliance rate, subjective questionnaire ratings on five-point Likert scales, and response time. Repeated-measures ANOVA was conducted to examine the main and interaction effects of emotion, action, and modality on persuasiveness and user perceptions.

RESULTS

Experiment 1 (Persuasive Actions in MR)

In the first experiment, the Proposed Method group showed higher compliance compared to the Bark-Only group. Participants in the Proposed Method group responded more frequently to the agent's prompts, particularly in the feeding and tidying tasks. Questionnaire results indicated that the Proposed Method group rated the agent as significantly more persuasive and clearer in intention. However, there was some variation depending on the emotion expressed: sadness was rated as persuasive and trustworthy, while anger tended to be perceived negatively. These results suggested that combining emotional expressions with behavioral actions increased persuasiveness and improved the clarity of the agent's intention.

Experiment 2 (Comparison of MR and Video)

In the second experiment, compliance rates were generally higher in the MR condition than in the video condition. Participants reported a stronger sense of social presence in MR, which was associated with higher ratings of persuasiveness. Repeated-measures ANOVA revealed significant main effects of modality, emotion, and action, as well as interaction effects between emotion and action. Sadness combined with pointing or guiding was rated as most persuasive, while anger combined with attention calling was rated least persuasive. Happiness was effective in some tasks such as tidying but less effective in discouraging smartphone use.

Response times were shorter in MR than in the video condition, indicating that participants understood the agent's intentions more quickly when the agent was co-present. Questionnaire ratings also showed that MR agents were perceived as more natural and trustworthy compared to video agents. Qualitative comments suggested that participants felt "the agent was really present in the room" in MR, whereas in video, some participants described the agent as "detached" or "less engaging."

Overall, the results of both experiments demonstrate that pet-type agents are more persuasive when emotional expressions are combined with clear behavioral actions, and that MR presentation enhances both persuasiveness and social presence compared to video.

DISCUSSION

The results of the two experiments provide new insights into the design of persuasive pet-type virtual agents. Experiment 1 demonstrated that combining emotional expressions with behavioral actions is more effective than simple signaling such as barking alone. This suggests that persuasive impact arises not merely from the presence of an agent but from the integration of expressive and communicative behaviors that clarify the agent's intention.

Experiment 2 further revealed that Mixed Reality presentation enhances persuasiveness compared to video. Participants reported higher social presence in MR, which in turn was associated with stronger compliance and more positive impressions of the agent. These findings are consistent with previous research showing that co-presence in MR fosters greater engagement and trust (Lystbæk et al., 2022) and also align with studies on proxemics in human–robot interaction, where users' acceptance of a robot's behavior depends on perceived social roles and appropriate interpersonal distance (Koay et al., 2014). This work extends those findings by demonstrating that persuasive outcomes are also improved.

The analysis of emotion–action combinations provides additional implications for design. Sadness combined with pointing or guiding was consistently rated as persuasive and trustworthy, indicating that empathy-driven cues are powerful in influencing user behavior. Happiness was effective in supportive contexts such as tidying tasks, while anger tended to be counterproductive. These findings emphasize the need for context-sensitive design: the same emotion may support persuasion in some situations but hinder it in others.

From a human-centered design perspective, the study demonstrates that pet-type agents offer a natural metaphor for persuasion. Because people are accustomed to interpreting animal behaviors as meaningful, leveraging such cues in virtual agents provides an intuitive and socially acceptable way to guide user behavior. Applications may extend to education, healthcare, and eldercare, where gentle and familiar forms of persuasion are preferable.

Nevertheless, several limitations should be noted. The participants were primarily university students, which limits the generalizability of the results. The set of scenarios was restricted to daily tasks in controlled experimental conditions, and future work should explore long-term interactions in real-world environments. Moreover, while the agent's actions were carefully designed, they were pre-scripted; adaptive behaviors based on real-time user responses may yield stronger persuasive effects.

In summary, this research shows that integrating emotional expressions and behavioral actions in pet-type agents significantly enhances persuasiveness, and that MR presentation provides additional benefits over traditional video. These insights contribute to the development of human-centered persuasive systems that can support positive behavior change in everyday contexts.

CONCLUSION

This paper presented two experiments investigating persuasive interactions with a pet-type virtual agent that combined emotional expressions and behavioral actions. Experiment 1 demonstrated that integrating emotions with behavioral cues was more persuasive than simple barking alone, highlighting the importance of expressive and communicative behaviors. Experiment 2 showed that Mixed Reality presentation enhanced persuasiveness compared to video, with participants reporting stronger social presence, higher compliance, and faster responses in MR.

Across both studies, sadness combined with guiding or pointing emerged as particularly effective, while anger was generally counterproductive. These findings underscore the need for context-sensitive design of persuasive agents and suggest that empathy-driven cues are especially powerful.

The contributions of this work are threefold. First, it demonstrates the persuasive potential of pet-type virtual agents as a natural metaphor for human–computer interaction. Second, it provides empirical evidence that Mixed Reality enhances persuasiveness by fostering social presence. Third, it offers design guidelines for combining emotions and actions in ways that are effective yet socially acceptable.

Future work should extend these findings by testing with more diverse populations, exploring long-term interactions in real-world contexts, and developing adaptive systems that respond dynamically to user behavior.

In conclusion, this research shows that pet-type agents in MR can support positive behavior change in everyday life, contributing to the development of human-centered persuasive technologies.

ACKNOWLEDGMENT

This work was supported by the New Energy and Industrial Technology Development Organization (NEDO) under Project JPNP23025.

REFERENCES

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human–computer relationships. ACM Transactions on Computer-Human Interaction, 12(2), 293–327. https://doi.org/10.1145/1067860.1067867

Bradshaw, J., & Rooney, N. (2016). Social and communication behaviour. In J. Serpell (Ed.), The domestic dog: Its evolution, behavior and interactions with people (pp. 133–159). Cambridge University Press. https://doi.org/10.1017/9781139161800.008

Breazeal, C. (2003). Emotion and sociable humanoid robots. International Journal of Human–Computer Studies, 59(1–2), 119–155. https://doi.org/10.1016/S1071-5819(03)00018-1

Cassell, J., Bickmore, T., Campbell, L., Vilhjálmsson, H., & Yan, H. (2000). Human conversation as a system framework: Designing embodied conversational agents. In J. Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), Embodied conversational agents (pp. 29–63). MIT Press. https://doi.org/10.7551/mitpress/2697.003.0004

Cassell, J., & Stone, M. (1999). Living hand to mouth: Psychological theories about speech and gesture in interactive dialogue systems. In AAAI 1999 Fall Symposium on Narrative Intelligence (pp. 34–42).

Fogg, B. J. (2002). Persuasive technology: Using computers to change what we think and do. Ubiquity, 2002(December), Article 2. https://doi.org/10.1145/764008.763957

Fujita, M. (2001). AIBO: Toward the era of digital creatures. The International Journal of Robotics Research, 20(10), 781–794. https://doi.org/10.1177/02783640122068092

Jones, C. M., & Deeming, A. (2007). Investigating emotional interaction with a robotic dog. In Proceedings of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI '07) (pp. 183–186). ACM. https://doi.org/10.1145/1324892.1324927

Kendon, L. (2004). Gesture: Visible action as utterance. Cambridge University Press.

Koay, K. L., Syrdal, D. S., Ashgari-Oskoei, M., Walters, M. L., & Dautenhahn, K. (2014). Social roles and baseline proxemic preferences for a domestic service robot. International Journal of Social Robotics, 6(4), 469–488. https://doi.org/10.1007/s12369-014-0232-4

Lystbæk, M. N., Rosenberg, P., Pfeuffer, K., Grønbæk, J. E., & Gellersen, H. (2022). Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality. Proceedings of the ACM on Human–Computer Interaction, 6(ETRA), Article 145, 18 pages. https://doi.org/10.1145/3530886

McNeill, D. (1992). Hand and mind: What gestures reveal about thought. University of Chicago Press.

Midden, C., McCalley, T., Ham, J., & Zaalberg, R. (2008). Using persuasive technology to encourage sustainable behavior. Journal of Applied Mechanics – Transactions of the ASME.

Miklósi, Á., Polgárdi, R., Topál, J., & Csányi, V. (1998). Use of experimenter-given cues in dogs. Animal Cognition, 1(2), 113–121. https://doi.org/10.1007/s100710050016

Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human–robot conversations: How robots might shape participant roles using gaze cues. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI '09) (pp. 61–68). ACM. https://doi.org/10.1145/1514095.1514109

Niewiadomski, R., & Pelachaud, C. (2007). Model of facial expressions management for an embodied conversational agent. In Proceedings of ACII ’07 (pp. 12–23). Springer. https://doi.org/10.1007/978-3-540-74889-2_2

Oinas-Kukkonen, H., & Harjumaa, M. (2018). Persuasive systems design: Key issues, process model and system features. In M. Howlett & I. Mukherjee (Eds.), Routledge handbook of policy design (pp. 87–105). Routledge.

Okada, M. (2023). *Weak robots: Relational-oriented approach to human well-being*. In H. Takeda (Ed.), *Systems design based on the benefits of inconvenience* (pp. 57–68). Springer. https://doi.org/10.1007/978-981-19-9588-0_6

Peng, X. B., Coumans, E., Zhang, T., Lee, T.-W., Tan, J., & Levine, S. (2020). Learning agile robotic locomotion skills by imitating animals. In Proceedings of RSS Workshops. arXiv:2004.00784

Raiola, G., Hoffman, E. M., Focchi, M., Tsagarakis, N., & Semini, C. (2020). A simple yet effective whole-body locomotion framework for quadruped robots. Frontiers in Robotics and AI, 7, Article 528473. https://doi.org/10.3389/frobt.2020.528473

Soproni, K., Miklósi, Á., Polgárdi, R., & Csányi, V. (2001). Comprehension of human communicative signs in pet dogs. Journal of Comparative Psychology, 115(2), 122–126. https://doi.org/10.1037/0735-7036.115.2.122

van der Borg, J., Schilder, M., Vinke, C., de Vries, H., & Matthijs, J. (2015). Dominance in domestic dogs: A quantitative analysis of its behavioral measures. PLOS ONE, 10(8), e0133978. https://doi.org/10.1371/journal.pone.0133978

Wada, K., & Shibata, T. (2007). Living with seal robots—Its sociopsychological and physiological influences on the elderly at a care house. IEEE Transactions on Robotics, 23(5), 972–980. https://doi.org/10.1109/TRO.2007.906261