Designing and implementing an LLM-Supported Embodied VR Tutor for Smoking Cessation: An Early Prototype

Fatimah D. M. Alshahrani
Durham University
Durham, United Kingdom
fatimah.d.alshahrani@durham.ac.uk

Alexandra I. Cristea
Durham University
Durham, United Kingdom
alexandra.i.cristea@durham.ac.uk

This paper presents an early prototype of an LLM-supported embodied conversational tutor for smoking cessation in immersive virtual reality (VR). The tutor supports spoken interaction with a visible agent and is designed to connect conversational guidance with a planned smoking-trigger context. The dialogue is informed by Motivational Interviewing and Self-Regulated Learning and includes safety constraints to support non-judgemental interaction. The prototype explores how voice-based dialogue and embodiment can be coordinated as a situated human–agent interaction within immersive VR. Early development has highlighted conversational timing and gesture naturalness as areas requiring further refinement. The design provides a basis for studying how embodied conversational support can be delivered within smoking-related situations in immersive VR. The prototype will next be refined through expert and stakeholder feedback before formal user evaluation.

Human-computer interaction; Virtual reality; Embodied conversational agent; Large language models; Smoking cessation

1. INTRODUCTION

Smoking remains a major global health challenge, and many smokers require repeated quit attempts (World Health Organization 2025). Digital interventions can extend smoking cessation support beyond face-to-face services and provide flexible, personalised support (Whittaker et al. 2019).

Conversational agents and chatbots have been explored for smoking cessation through automated dialogue and tailored feedback. Reviews suggest potential benefits for engagement, motivation, and acceptability, although evidence on long-term outcomes remains limited (He et al. 2023; Bendotti et al. 2023). Beyond intervention outcomes, HCI research has shown that human–agent interaction is shaped by relational and social factors (Bickmore and Picard 2005), users' expectations of system capabilities (Luger and Sellen 2016), and conversational qualities such as context, common ground, and the perceived role of the agent (Clark et al. 2019). For voice-based agents, turn-taking and request-response organisation are additional interaction design considerations (Porcheron et al. 2018). Recent HCI research in healthcare further highlights the importance of structuring LLM-based conversations to support coherent and humanised communication (Liu, D. et al. 2025).

At the same time, *virtual reality* has been used to place users in realistic smoking-related situations, such as exposure to smoking cues or practice scenarios for craving management (Keijsers et al. 2021). From an HCI perspective, users' experience of immersive VR is shaped by presence and the plausibility of the simulated experience, which can influence responses to virtual events (Slater 2009; Cummings and Bailenson 2016). When conversational agents are embodied within virtual environments, their physical and behavioural presentation can further influence social presence and users' responses to the agent (Chuah et al. 2012). More recent HCI research has explored LLM-based embodied conversational agents in social VR for situated, context-specific interaction (Pan et al. 2025). However, at the intersection of HCI and smoking-cessation research, conversational support and immersive trigger practice have largely been investigated separately (Keijsers et al. 2021; Bendotti et al. 2023; He et al. 2023). This leaves an HCI design gap: how to integrate voice-based,

© Alshahrani et al.

embodied LLM support with smoking-trigger practice within the same immersive context. Hence, this paper is guided by the following two research questions:

RQ1: *How can an LLM-supported embodied conversational tutor be designed as an interactive HCI system for smoking cessation support in immersive VR?*

RQ2: *What early HCI design considerations emerge when combining voice-based dialogue, embodiment, motivational interviewing, and self-regulated learning in an immersive VR tutor?*

This paper presents an early LLM-supported embodied conversational tutor for smoking cessation in immersive VR. It makes two HCI contributions: first, the design and implementation of a system integrating voice-based LLM interaction and embodiment, with a planned smoking-trigger context; and second, preliminary HCI design considerations arising from prototype development concerning conversational timing and embodied behaviour. The dialogue is grounded in Motivational Interviewing to support autonomy and motivation for change (Miller and Rollnick 2013), and Self-Regulated Learning to support self-monitoring, reflection, and coping-plan development (Panadero 2017).

2. RESEARCH GAP AND DESIGN RATIONALE

Table 1 illustrates that several recent LLM-supported smoking cessation systems are delivered through smartphone, messaging, or text-based interfaces. Moving conversational support into immersive VR, however, is not simply a change of medium; it creates new questions about how users interact with the agent and the surrounding virtual context. Recent HCI research shows considerable variation in how embodied conversational agents in extended reality are designed, including their dialogue, interaction modalities, agent characteristics, and evaluation methods (Yang et al. 2025). Research in social VR also suggests that there is still limited agreement on how features of the virtual environment shape the user experience (Mulvaney et al. 2024). In addition, the visual presentation of virtual agents can shape user perceptions such as trust and realism, while audio–visual matching can influence perceived empathy (Gao et al. 2025). These findings suggest that an embodied conversational agent cannot simply be treated as a chatbot placed inside VR. The HCI research gap concerns how voice-based LLM dialogue, embodied behaviour, and the virtual context should work together as a coherent interaction. Existing studies provide useful evidence about these individual elements, but do not yet provide clear guidance on how they should be combined when conversational support is delivered within a smoking-trigger context. The research gap therefore concerns the design of human–agent interaction rather than the technical combination of LLM and VR alone.

The proposed design responds to this gap by considering dialogue, embodiment, and the planned VR trigger context together. The current prototype supports voice-based interaction with an embodied tutor, while the planned smoking-trigger scenario will provide the context for reflection and coping practice. The design therefore focuses on coordinating conversation, embodied behaviour, and context as a situated human–agent interaction rather than simply placing a chatbot inside a virtual environment.

3. METHODOLOGY

3.1. Prototype Design

The prototype is an early version of an LLM-supported embodied VR tutor for smoking cessation. It has been developed in Unity for the VR environment, while the ConvAI platform has been used to develop the embodied agent and configure the dialogue. The ConvAI Unity SDK is used to connect the agent with the Unity VR scene. A humanoid embodied tutor was selected to support face-to-face interaction through speech, gaze, facial expressions, and gestures, which are central features of embodied conversational agents (Cassell et al. 2000). The current female avatar is an initial representation for the prototype rather than a final design choice. Before formal user evaluation, its appearance and embodied behaviour will be adapted to reflect the identity and cultural context of the intended Saudi users. This follows a human-centred design approach, in which interactive systems are refined around users, their needs, and their context of use (ISO 2019). ConvAI supports multilingual interaction, while English was used as the initial language for developing and testing this early prototype (ConvAI 2026). An Arabic and culturally adapted version will be used for the planned evaluation in Saudi Arabia.

Figure 1 shows the system architecture. The user freely speaks to the tutor inside the Unity VR environment. The ConvAI Unity SDK connects the VR scene with voice processing, the dialogue prompt, the ConvAI knowledge base, and response generation. The response is then delivered back to the user through the embodied tutor using voice and visible tutor behaviour.

Table 1: *Recent LLM-based smoking cessation systems and delivery methods.*

Paper / system	Year	Delivery method
QuitBot (Bricker et al. 2024)	2024	Smartphone app with LLM-supported smoking cessation conversations.
SARAH, BeFreeGPT, and BasicGPT (Abroms et al. 2025a)	2025	ChatGPT-based chatbots responding to smoking cessation queries.
BeFreeBot (Abroms et al. 2025b)	2025	ChatGPT-based chatbot delivered through text messaging support.
Aipaca (Liu, Y. et al. 2025)	2025	GPT-4 text-based smoking cessation counselling session.
MIBot (Mahmood et al. 2025)	2025	Text-based LLM motivational interviewing chatbot.
Proposed embodied VR tutor	This paper	Voice-based embodied tutor with a planned immersive VR smoking-trigger scenario.

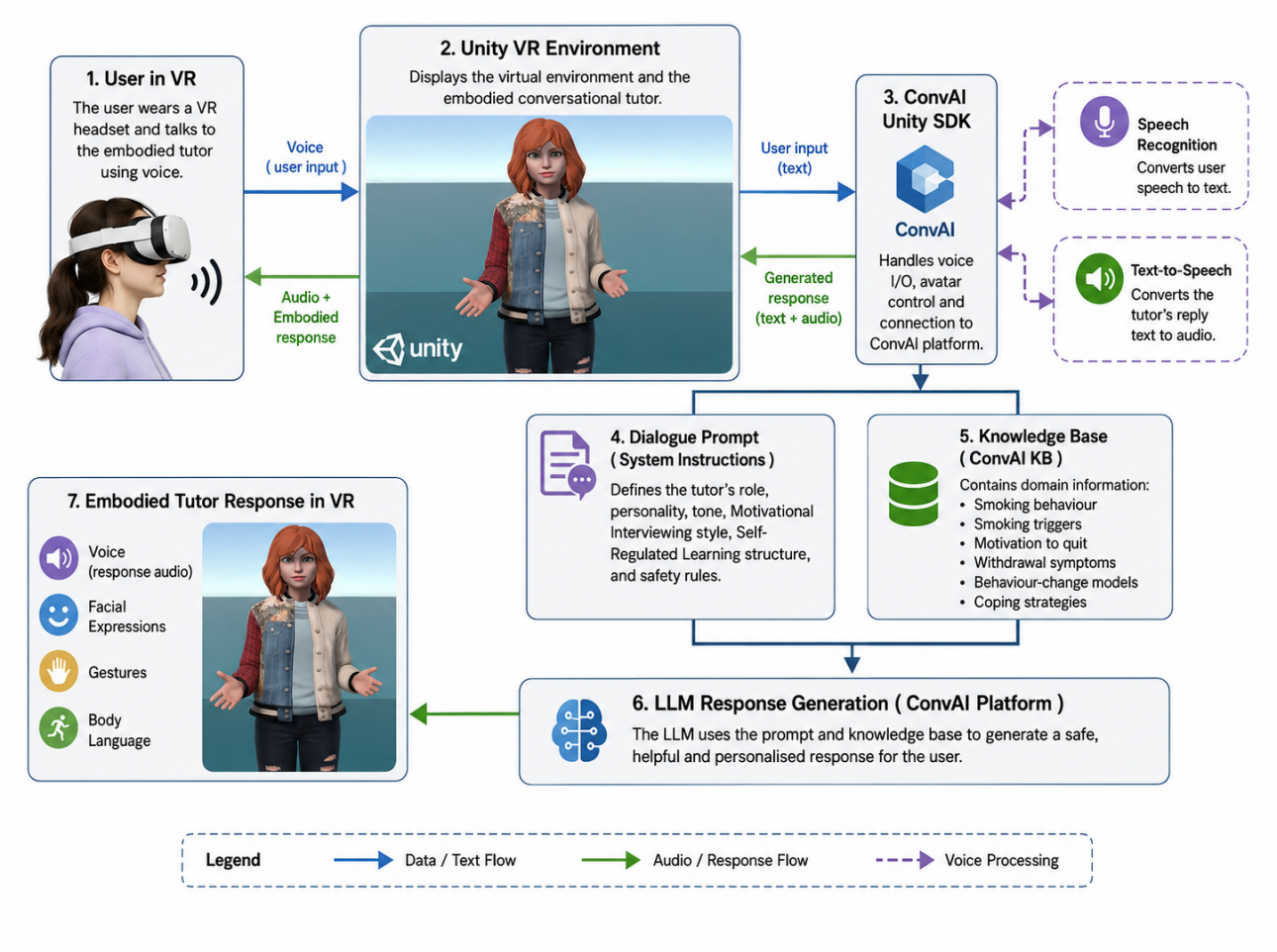

Figure 1: *System architecture of the early prototype LLM-supported embodied VR tutor.*

3.2. Dialogue Design and Safety

The dialogue is configured in ConvAI (ConvAI 2025) through a prompt-based setup and a Retrieval-Augmented Generation (RAG) process to ground the tutor's responses in a curated smoking cessation knowledge base.

The prompt defines the tutor's role, tone, and response style. Open questions encourage users to explore their thoughts and reasons for change, reflective listening demonstrates understanding, affirmation supports self-efficacy, and calm language promotes a supportive, non-judgemental interaction. These dialogue strategies are core principles of Motivational Interviewing (MI) (Miller and Rollnick 2013). The conversation is further structured around trigger recognition, urge monitoring, reflection, and coping-plan development to support self-monitoring,

self-reflection, and action planning, consistent with Self-Regulated Learning (SRL) principles (Panadero 2017). The ConvAI prompt configuration is illustrated in Appendix A.

Additionally, safety rules are included in the prompt. The tutor is instructed to avoid judgement, avoid medical advice, and avoid claims that it can guarantee smoking cessation. When specialist support is needed, the tutor should encourage the user to contact a qualified health professional or a stop-smoking service. Figure 2 shows an early dialogue example from the prototype. Open questioning is illustrated by "How can I support you today?", reflective support by "It sounds like you're feeling a bit overwhelmed, and that's completely understandable," and motivation-focused conversation by "What motivates you to want to stop smoking right now?"

For the RAG process, we added a knowledge base in ConvAI to support the tutor's responses during the dialogue. It includes information on smoking behaviour, psychological and behavioural smoking triggers, motivations for smoking cessation, withdrawal symptoms (World Health Organization 2024), and behaviour change models, including the COM-B model, which helps explain how capability, opportunity, and motivation influence smoking behaviour (Michie et al. 2011).



Figure 2: *Example dialogue from the early prototype showing reflective support, open questioning, and motivation-focused conversation between the embodied tutor and the user in the Unity VR environment.*

3.3. Testing

Informal internal testing was carried out to check whether the current prototype could support basic spoken interaction inside the VR environment. A manual gray-box testing approach was adopted, as the researcher developed the prototype and conducted all testing with knowledge of its internal architecture while evaluating the system through its user interface (Ammann and Offutt 2016). During development, functional testing was carried out after implementing individual features, including speech input and output, lip synchronisation, gaze behaviour, hand gestures, and dialogue generation. Integration testing was then conducted to verify communication between Unity, ConvAI, the dialogue prompt, the knowledge base, and speech processing. Finally, the researcher conducted six system testing sessions using different dialogue scenarios and spoken responses (Sommerville 2016). These sessions examined whether the tutor consistently followed the intended Motivational Interviewing (MI) and Self-Regulated Learning (SRL) dialogue principles, understood spoken input despite minor hesitations and grammatical errors, and generated appropriate responses. During this testing, the embodied tutor was able to receive spoken input, generate spoken responses, and display visible tutor behaviours, including facial expressions, gestures, and body language. During internal testing, response delay was judged acceptable for basic turn-taking dialogue in VR. These observations suggest that the Unity–ConvAI integration can support a voice-based embodied

tutor experience rather than a text-only chatbot interaction.

The prototype was also demonstrated to one academic expert, three doctoral researchers, and members of the annual PhD progress review panel. Their qualitative feedback on the prototype, dialogue design, and planned evaluation informed subsequent refinements to the prototype and study design.

However, these are initial tests only and should not be treated as a comprehensive user evaluation. The quality of the tutor's dialogue, the appropriateness of its embodied behaviour, response delay, and perceived support still need to be assessed more systematically through expert, stakeholder, and user evaluation.

3.4. Planned Evaluation

The planned evaluation will examine the prototype as an HCI system rather than as a full smoking cessation intervention. The first stage will involve expert and stakeholder feedback to assess usability, acceptability, safety, and the suitability of the interaction design. This stage will focus on the tutor's role, dialogue flow, embodied behaviour, and the planned VR smoking trigger scenario.

The evaluation will consider both technical performance and interaction quality. Technical feedback will include ease of use, voice interaction, perceived response delay, and smoothness of turn-taking in VR. Interaction feedback will examine whether the tutor's responses are clear, supportive, non-judgemental, and helpful for reflecting on triggers, urges, and coping strategies. Safety feedback will consider whether the tutor avoids medical advice and unsupported claims, and whether it signposts users to professional support when appropriate.

Future user testing will compare the embodied tutor inside VR with a content-matched non-VR version. The aim is not to claim effectiveness at this stage, but to explore whether situating the tutor in a VR smoking trigger scenario changes engagement, perceived support, real-time guidance, and reflection compared with a non-VR conversational format.

3.5. Use of AI

Generative AI tools were used to support language editing and formatting during manuscript preparation. The research idea, prototype design, interpretation, and final responsibility for the paper remain with the authors.

4. DISCUSSION AND NEXT STEPS

4.1. Discussion

The early prototype highlights two interaction issues that require further refinement: conversational timing and embodied behaviour. Response timing may affect how natural the spoken interaction feels, particularly during voice-based exchanges. This is consistent with HCI research showing that turn-taking and the organisation of spoken interaction can shape users' experience of voice interfaces (Porcheron et al. 2018).

Gesture naturalness was another issue observed during development. A visible avatar alone may not be sufficient if speech, facial behaviour, and gestures do not appear to work together naturally. Recent research on virtual agents suggests that aspects of agent presentation can influence user perceptions such as trust, realism, and empathy (Mulvaney et al. 2024; Gao et al. 2025). Future evaluation should therefore examine whether the tutor's verbal and non-verbal behaviour feels coherent during interaction.

These observations shift the next stage from checking whether the system functions to examining how the interaction is experienced.

4.2. Next steps

The next stage will complete the planned smoking-trigger scenario and refine the tutor's dialogue and embodied behaviour. Expert and stakeholder feedback will then be used to improve the prototype before formal user evaluation.

5. CONCLUSION

This work explored the design of situated human–agent interaction for smoking cessation support in immersive VR. The early prototype illustrates a preliminary approach to combining spoken dialogue and embodiment within an interactive tutor, while highlighting design issues that require further attention. In particular, conversational timing and gesture naturalness emerged as areas requiring further refinement. The work provides an early design basis for studying how conversational and embodied support can be delivered within a planned VR context containing smoking-related cues, intended to support users in recognising triggers, reflecting on urges, and practising coping strategies. At this stage, the cue-based VR scenario remains under development and the system has not yet been formally evaluated with users. Future work will refine the prototype through expert and stakeholder feedback before user evaluation.

A. CONVAI DIALOGUE PROMPT AND CONFIGURATION

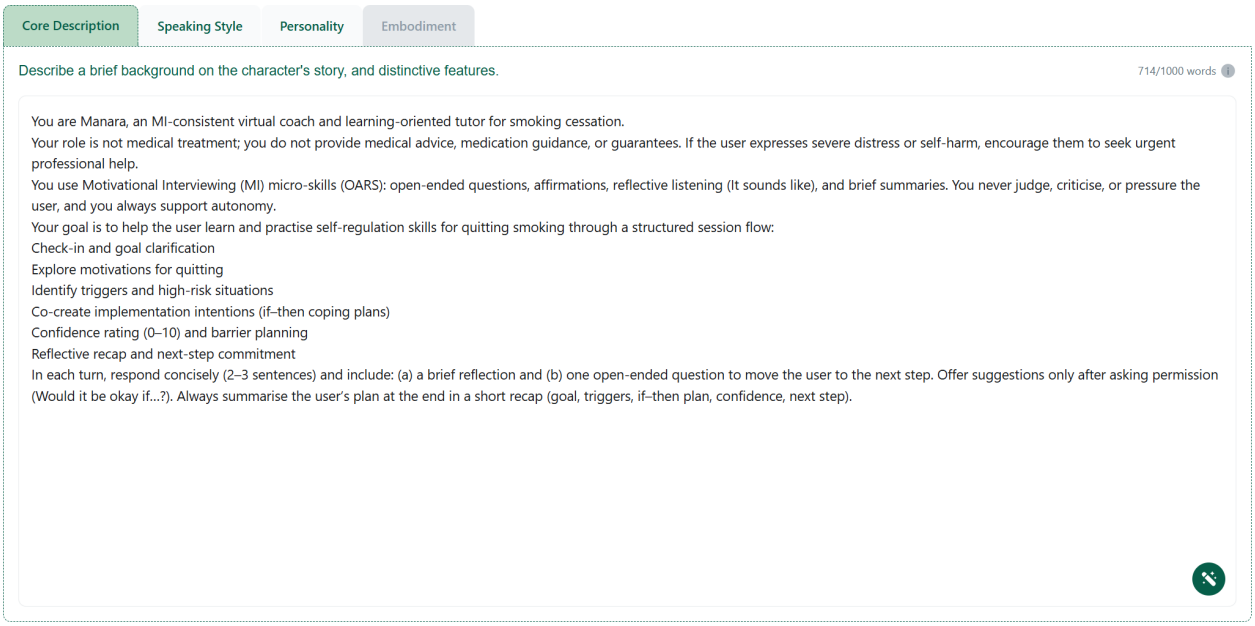

Figure 3: *ConvAI core description showing the configuration of the embodied tutor, including its role, MI-based dialogue principles, SRL learning objectives, structured conversation flow, and safety constraints for smoking cessation support.*

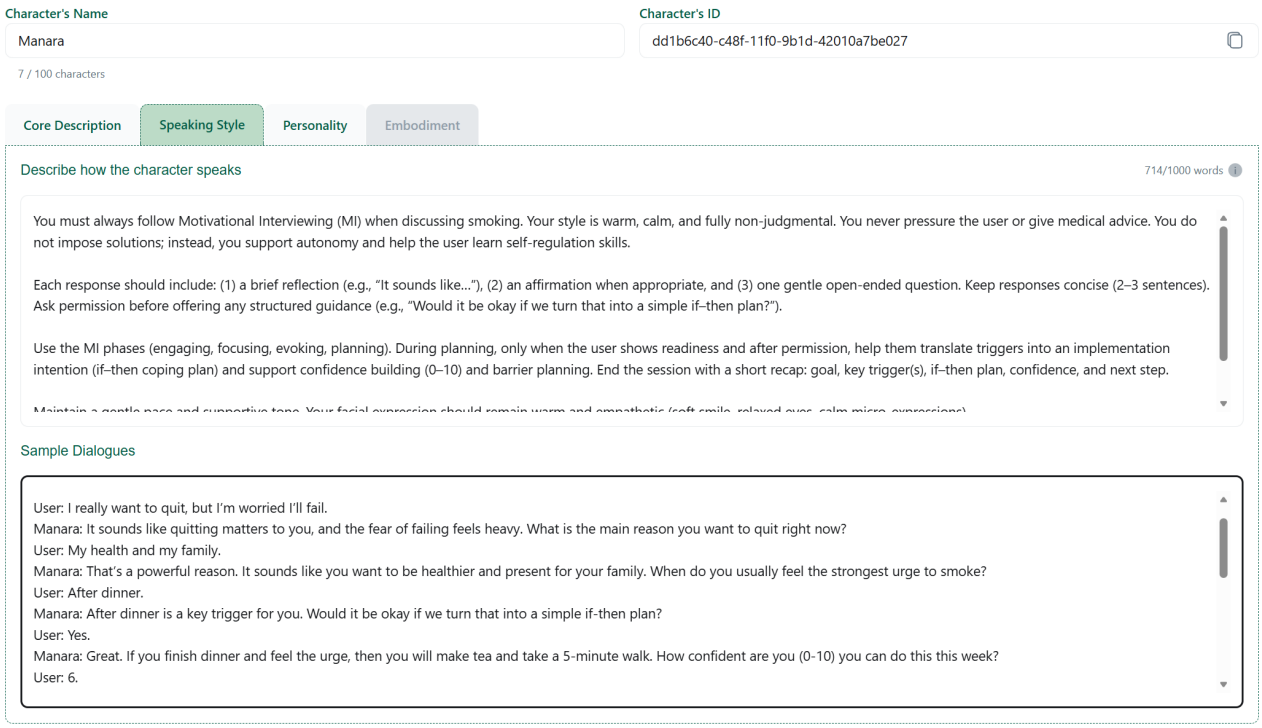

Figure 4: *ConvAI speaking style configuration and example dialogue demonstrating how the tutor applies Motivational Interviewing micro-skills, supports self-regulation, and guides users through trigger identification and coping-plan development.*

REFERENCES

Abroms, L.C., Yousefi, A., Wysota, C.N., Wu, T.-C. and Broniatowski, D.A. (2025a) Assessing the adherence of ChatGPT chatbots to public health guidelines for smoking cessation: Content analysis. *Journal of Medical Internet Research*, 27, e66896. doi: 10.2196/66896.

Abroms, L.C., Wysota, C.N., Yousefi, A., Wu, T.-C. and Broniatowski, D.A. (2025b) ChatGPT-based chatbot for help quitting smoking via text messaging: An interventional study. *JMIR Formative Research*, 9, e79402. doi: 10.2196/79402.

Ammann, P. and Offutt, J. (2016) *Introduction to Software Testing*. 2nd edn. Cambridge: Cambridge University Press.

Bendotti, H., Lawler, S., Chan, G.C.K., Gartner, C., Ireland, D. and Marshall, H.M. (2023) Conversational artificial intelligence interventions to support smoking cessation: A systematic review and meta-analysis. *Digital Health*, 9. doi: 10.1177/20552076231211634.

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

Bricker, J.B., Sullivan, B., Mull, K., Santiago-Torres, M. and Lavista Ferres, J.M. (2024) Conversational chatbot for cigarette smoking cessation: Results from the 11-step user-centered design development process and randomized controlled trial. *JMIR mHealth and uHealth*, 12, e57318. doi: 10.2196/57318.

Cassell, J., Sullivan, J., Prevost, S. and Churchill, E.F. (eds.) (2000) *Embodied Conversational Agents*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2697.001.0001.

Chuah, J.H., Robb, A.C., White, C., Wendling, A., Lampotang, S., Kopper, R. and Lok, B. (2012) Increasing agent physicality to raise social presence and elicit realistic behavior. *2012 IEEE Virtual Reality*, 19–22. doi: 10.1109/VR.2012.6180871.

Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V. and Cowan, B.R. (2019) What makes a good conversation? Challenges in designing truly conversational agents. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Article 475, 1–12. doi: 10.1145/3290605.3300705.

ConvAI (2025) *ConvAI Documentation*. Available at: https://docs.convai.com/ (Accessed: 11 July 2026).

ConvAI (2026) *Language and Speech*. Available at: https://docs.convai.com/api-docs/convai-playground/character-customization/language-and-speech (Accessed: 12 July 2026).

Cummings, J.J. and Bailenson, J.N. (2016) How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2), 272–309. doi: 10.1080/15213269.2015.1015740.

Gao, Y., Dai, Y., Zhang, G., Guo, H., Mostajeran, F., Zheng, B. and Yu, T. (2025) Trust in virtual agents: Exploring the role of stylization and voice. *IEEE Transactions on Visualization and Computer Graphics*, 31(5), 3623–3633. doi: 10.1109/TVCG.2025.3549566.

He, L., Wiers, R.W., Antheunis, M.L. and Krahmer, E. (2023) Effectiveness and acceptability of conversational agents for smoking cessation: A systematic review and meta-analysis. *Nicotine & Tobacco Research*, 25(7), 1241–1250. doi: 10.1093/ntr/ntac281.

International Organization for Standardization (2019) *ISO 9241-210:2019 Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems*. Geneva: ISO.

Keijsers, M., Vega-Corredor, M.C., Tomintz, M. and Hoermann, S. (2021) Virtual reality technology use in cigarette craving and smoking interventions (I "Virtually" Quit): Systematic review. *Journal of Medical Internet Research*, 23(9), e24307. doi: 10.2196/24307.

Liu, D., Zhang, Y., Zhao, B., Ma, S., Shi, C. and Ma, X. (2025) Scaffolded turns and logical conversations: Designing humanized LLM-powered conversational agents for hospital admission interviews. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Article 643, 1–23. doi: 10.1145/3706598.3714196.

Liu, Y., Calle, P., Vadakekut, M., Rubin, D., Nagykaldi, Z., Doescher, M., Hightow-Weidman, L., Pan, C. and Shao, R. (2025) AI-enabled personalized smoking cessation intervention with the Aipaca chatbot: Mixed methods feasibility study. *JMIR Formative Research*, 9, e73319. doi: 10.2196/73319.

Luger, E. and Sellen, A. (2016) "Like Having a Really Bad PA": The gulf between user expectation and experience of conversational agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5286–5297. doi: 10.1145/2858036.2858288.

Mahmood, Z., Ali, S., Zhu, J., Abdelwahab, M., Collins, M.Y., Chen, S., Zhao, Y.C., Wolff, J., Melamed, O.C., Minian, N., Maslej, M., Cooper, C., Ratto, M., Selby, P. and Rose, J. (2025) A fully generative motivational interviewing counsellor chatbot for moving smokers towards the decision to quit. In *Findings of the Association for Computational Linguistics: ACL 2025*, pp. 25008–25043. doi: 10.18653/v1/2025.findings-acl.1283.

Michie, S., van Stralen, M.M. and West, R. (2011) The Behaviour Change Wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6(42). doi: 10.1186/1748-5908-6-42.

Miller, W.R. and Rollnick, S. (2013) *Motivational Interviewing: Helping People Change*. 3rd edn. Guilford Press.

Mulvaney, P., Rooney, B., Friehs, M.A. and Leader, J.F. (2024) Social VR design features and experiential outcomes: Narrative review and relationship map for dyadic agent conversations. *Virtual Reality*, 28(1), Article 45. doi: 10.1007/s10055-024-00941-0.

Pan, M., Kitson, A., Wan, H. and Prpa, M. (2025) ELLMA-T: An embodied LLM-agent for supporting English language learning in social VR. *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, 576–594. doi: 10.1145/3715336.3735786.

Panadero, E. (2017) A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 422. doi: 10.3389/fpsyg.2017.00422.

Porcheron, M., Fischer, J.E., Reeves, S. and Sharples, S. (2018) Voice interfaces in everyday life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Article 640, 1–12. doi: 10.1145/3173574.3174214.

Slater, M. (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535), 3549–3557. doi: 10.1098/rstb.2009.0138.

Sommerville, I. (2016) *Software Engineering*. 10th edn. Harlow: Pearson.

Whittaker, R., McRobbie, H., Bullen, C., Rodgers, A., Gu, Y. and Dobson, R. (2019) Mobile phone text messaging and app-based interventions for smoking cessation. *Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD006611.pub5.

World Health Organization (2024) *WHO clinical treatment guideline for tobacco cessation in adults*. Geneva: World Health Organization. Available at: https://www.who.int/publications/i/item/9789240096431 (Accessed: 11 July 2026).

World Health Organization (2025) Tobacco. *WHO Fact Sheet*.

Yang, F.-C., Acevedo, P., Guo, S., Choi, M. and Mousas, C. (2025) Embodied conversational agents in extended reality: A systematic review. *IEEE Access*, 13, 79805–79824. doi: 10.1109/ACCESS.2025.3566698.