Human-Computer Interaction & Emerging Technologies, Vol. 233, 2026, 456–465
https://doi.org/10.54941/ahfe1007545

Co-Designing an Avatar-Based Agent for Cybersecurity Training in VR for Personnel in Critical Infrastructure Sectors

Vanessa Roberts, Tiia Sõmer, and Rain Ottis

Tallinn University of Technology, Tallinn, 19086, Estonia

ABSTRACT

This paper provides an overview of how to design an avatar-based agent virtual reality (VR) for cybersecurity training for the transport and water sector personnel using the human-centered design approach. An agent learner interactions (EnALI) framework in survey form has been used to gather input from a target group consisting of experts-in-training and specialists currently working in the field. Content and thematic analysis were used to identify what the avatar should be able to do, how it should be perceived and the reasons behind it. The data gathered has been analyzed to produce an agent persona and use cases. The outputs highlight the need for a human-like pedagogical and facilitator agent that would support training participants to meet their learning goals. The findings of this study provide valuable insights for researchers and developers as to the implementation of avatar-based agents in VR environments for cybersecurity training.

Keywords: Avatar-based agent, Cybersecurity training, Human-centered design, Agent learner interactions framework

INTRODUCTION

ENISA´s Cybersecurity Incident Reporting and Analysis System ("Incident reporting," 2023) shows that 7% of 540 reported incidents from 2020–2024 in the transport area are due to human errors. Human errors consist of issues related to policies and procedures, falling victim to phishing attacks, and malware and viruses. For the same period, 11% of the 141 incidents reported in the drinking water supply and distribution area are caused by human errors noting the impact of malware on industrial systems. The actual number of incidents caused by human error may be considerably higher due to underreporting emanating from lack of root cause analysis, monitoring, and reporting due to concern over loss of reputation. Therefore, training about the causes for incidents has been identified as one of the most important aspects (ENISA, 2024). Conducting this training in virtual environments, where scenarios are simulated, is a safe and engaging means to entice users to interact with the training setting (ECSO WG5, 2020).

To further motivate the case for innovative cybersecurity training approaches, the EU has noted a lack of cybersecurity specialists, including educators (ECSO WG5, 2017), (ENISA, 2021), (ENISA, 2024), to be able to train future experts to protect critical infrastructure.

Received February 6, 2026; Revised April 4, 2026; Accepted April 18, 2026; Available online July 20, 2026

Agents are a nascent technology in filling a specific role in the learning process (Adinolf et al., 2020). Currently, there is a sparsity of research on agents as avatars in VR for cybersecurity training. The key contribution of this paper is the application of the Human-Centered Design (HCD) approach for the design of agent-based avatars in VR. This paper aims to design an agent based on the target group´s needs. In this paper, we focus on the following critical research questions (RQ):

RQ1. What qualities and functionalities are important when developing an avatar-based agent in the context of cybersecurity training in VR?

RQ2. What are the underlying reasons that make those functionalities important?

BACKGROUND AND RELATED WORK ON DESIGNING AGENTS

Avatar-based agents (ABA), also referred to as pedagogical agents (PA), conversational agents (CA) or embodied conversational agents (ECAs) — are computer-generated 2D or 3D characters designed to simulate aspects of human interaction in virtual environments (Martha and Santoso, 2019), (Bente et al., 2023).

Traditionally, an avatar is a graphical presentation through which the user can interact with the (simulated) environment. Most of the research focuses on avatar customization in games for entertainment. In this case, an avatar is not the representation of the user itself as often depicted in games, but an assisting entity for the user who is a training participant. The avatar is designed to enhance learning, meaning, it exists in a specific context (Taylor, 2000). A context-specific avatar for cybersecurity is designed to support diverse cybersecurity training scenarios.

The potential of avatars was recognized in the field of education more than a decade ago, mainly in offering learning support. According to Baylor (2011), the appearance and message delivery (e.g., voice, non-verbal communication) together with the motivational message and dialogue are key design considerations for a motivational agent. However, describing the process of designing agents as avatars for a VR environment in the context of cybersecurity training context remained relatively unexplored, with the notable exception of (Adinolf et al., 2019). In the process of designing agent-based VR applications for cybersecurity training, their work found that the partner agent is not needed to be a human representation as it can also take the form of an anthropomorphic robot or an animal.

Research exists which explores various HCD aspects. However, there is a sparsity of research which uses HCD to design agents. Through engaging with a human-centered process in the design, the avatar interactions are designed to be inherently human-centered (Capel and Brereton, 2023). Their article has influenced the application of human-centered methods to a wide variety of contexts, such as health and medical domains.

According to ISO 9241-210:2019 (ISO - International Organization for Standardization, 2019), the principles of human-centered design entail understanding of users, tasks and environments as the users are involved in the design and development process, but also later in the evaluation

phase. The HCD process consists of the following activities with example outputs:

1. Understanding and specifying the context of use, where one of the outputs according to the HCD process is a persona.
2. Specifying the user requirements, where one of the outputs is that user needs and requirements have been identified and described using the use case format.
3. Producing design solutions to meet these requirements with a prototype as an output.
4. Evaluating the designs against requirements with an output of a report.

Another aspect of designing agents is having frameworks guiding the work. Baylor (Baylor, 2011) notes that there are very few agent-specific guidelines available. One of the very few existing frameworks which is based on academic research and is suitable for gathering input from the noted specialists in the transport and water sector, is the Enhancing Agent Learner Interactions (EnALI) (Veletsianos, George et al., 2009). This approach is grounded on three major theories: 1) socio-cultural notions of learning, 2) cooperative learning and 3) conflict theory. They have proposed a three-tier framework of 15 research-based guidelines focusing on interaction, message and agent characteristics – all relevant when designing an agent.

This current paper contributes towards designing and studying agent-based VR training environments by applying the HCD process and the EnALI framework. For the purposes of this article, we focus on the first three steps of the HCD process. The last step "Evaluating the designs against requirements" is left out of scope as the visual presented is not a fully interactive prototype.

METHODS

This study employs a mixed-methods approach. The data collected from the co-design workshops uses the EnALI framework, where 15 guidelines were turned into statements that survey participants could assess using a Likert scale (1 = strongly disagree, to 5 = strongly agree). The quantitative data derived from the Likert scale in the Google Forms is integrated into the qualitative analysis.

The respondents were able to elaborate on guidelines/statements by using the free-text input area, then for qualitative data analysis, functional requirements and possible agent qualities were extracted. Content analysis was conducted to identify patterns. In addition, a thematic analysis was conducted for a deeper interpretation as it can also be used with textual data from qualitative surveys (Terry et al., 2017).

Participants

The pilot study was conducted with vocational school and university students studying information technology. Out of 37 participants, 26 participants provided input at the summer camp in 2025 in Estonia.

For the main study, over the course of the following months, the target group consisted of specialists working in the transport sector (waterways and maritime) and water sector (drinking and wastewater area). In addition, experts-in-training studying at various levels of education (vocational education, applied sciences, bachelor, master and PhD students). During several co-design workshops, 114 participants answered the questionnaire. As participants were able to opt out and submit an empty form, a total number of 90 participants gave specific input to help design the avatar.

Materials

The materials for this workshop included:

- a short lecture and an introduction via MS PowerPoint
- a survey in Google Forms consisting of the informed consent and the statements for data gathering.

Procedure

The co-design workshops were held in hybrid mode (online and offline) and lasted for one hour. The workshop started with a project description that also included a photo of the VR room and MS Clippy as an inspiration for the avatar. A short 20-minute lecture on emerging technologies was conducted to incentivize participants. The goal of the workshop was introduced which highlighted the consent form and explained the logic of the questionnaire. The participants were asked to complete the questionnaire by themselves without using any additional tools. The workshop concluded with information about the next steps of the development process.

DATA ANALYSIS

Quantitative Analysis

The original EnALI framework consisting of 3 parts was used in the questionnaire for data collection. The respondents´ level of agreement, indicating also the relevance to the target group, are described below (see Table 1).

1. User interaction (agent-user interaction). The respondents mostly strongly agree on the importance of responsiveness and reactiveness by the agent to the user´s requests for additional and/or expanded information.
2. Message (agent communication). The respondents mostly strongly agree that the agent should craft a message appropriate to the receiver's abilities, experiences and frame of reference. Furthermore, it is mostly strongly agreed that the agent should craft messages that are complete and specific.

3. Agent characteristics. The target group strongly agrees on the importance of the agent establishing credibility and trustworthiness, the agent being polite and positive (e.g., encouraging, motivating) and the agent using a visual representation appropriate to content.

Table 1: The respondents´ level of agreement with EnALI framework aspects (N = 90).

Framework Tier and Design Guideline/Statement	Likert Scale Average (Standard Deviation)
Interaction	4.2 (0.8)
1.1 Being responsive and reactive to requests for additional and/or expanded information.	2.5 (1.2)
1.2 Being redundant.	3.9 (1.0)
1.3 Asking for formative and summative feedback.	3.9 (1.0)
1.4 Maintaining an appropriate balance between on- and off-task communications.	
Message	4.3 (0.9)
2.1 Making the message appropriate to the receiver's abilities, experiences, and frame of reference.	3.8 (1.3)
2.2 Using congruent verbal and non-verbal messages.	4.0 (1.0)
2.3 Clearly owning the message.	4.5 (0.8)
2.4 Making messages complete and specific.	3.6 (1.1)
2.5 Using descriptive, non-evaluative comments.	3.4 (1.3)
2.6 Describing feelings by name, action, or figure of speech.	
Agent´s characteristics	4.3 (1.0)
3.1 Establishing credibility and trustworthiness.	3.7 (1.1)
3.2 Establishing role and relationship to user/task.	4.4 (1.0)
3.3 Being polite and positive (e.g., encouraging, motivating).	3.5 (1.1)
3.4 Being expressive (e.g. exhibiting verbal cues in speech).	4.3 (1.0)
3.5 Using a visual representation appropriate to content.	

QUALITATIVE ANALYSIS

Based on the several co-design workshops organized, a context analysis was conducted by extracting information from the survey responses.

To avoid misinterpretation, the authors aimed at remaining as true as possible to the initial answer given by the respondent. For example, during the analysis, the focus was put on what the agent must/should do or needs to do and be like (perception aspects) as this showed a strong indication of having a specific preference by the respondent. In addition, unclear free text area answers were classified as "not clear enough to extract a functionality/adjective". Empty or uninformative free text area content was classified as "not informative".

One free text area answer could contain zero or more functionalities and adjectives. The text coding was conducted in a systematic way. The content that did contain functionalities or adjectives was further classified into categories of the Figma template "Chatbot or Voice Persona" (Caio Calado, 2022) to structure the data. The categories remaining after consolidating

similar ones are the following: personality traits, tone, behavior, capabilities. The coding schema is presented in Table 2.

Table 2: The coding schema for content analysis.

Category in Figma Template	Description Based on Cambridge´s Dictionary	Example From the Survey Free Text Comment Area
Personality traits	Characteristics describing an agent	"Virtual assistant should be responsive and helpful..."
Tone	A quality in the voice of an agent	"…being expressive in a positive tone"
Behavior	The way an agent behaves (including what it should avoid doing, how it reacts to success and failure considering what is it designed to be motivated by)	"Helps keep up spirit and motivation" "It shouldn't get in the way when it's not needed."
Capabilities	The agent´s ability to do something	"Agent should be on standby if not used actively"

Based on the information presented in Table 2, the output of personality traits, tone and behavior led to the creation of the persona. The examples in the third column are the input given by respondents. Considering the various ways of data interpretation, the examples may not be 100% applicable or may contain controversies.

The capabilities were considered as the agents' possible functionalities, and they were put into the unified modelling language (UML) use case diagram format focusing on what the agent should be able to do. Regarding the capabilities as agent´s possible functionalities, a thematic analysis was conducted to identify the underlying reasons that make those functionalities important (see Table 3).

Table 3: The coding schema for thematic analysis.

Example From the Survey Free Text Comment Area	Theme	Theme Description
"Agent should be on standby if not used actively"	User in control	Describing why the user prefers to be in control when using the agent
"Getting feedback is a must have in order to develop as a student."	Benefits for users	Describing why the agent needs to provide certain benefits for the user
"Messages should be specific that you can interpret the answers only one way and avoid misunderstandings"	Communication towards user	Describing why the agent´s communication needs to be in a certain way

RESULTS

Agent Persona

Perception aspects, meaning, how the agent should be perceived, were identified based on the survey data. The data analysis revealed duplicated functionalities which were consolidated. Most important perception aspects

were determined based on the number of mentions. The results were visualized with the help of a Figma agent persona template (see Figure 1).

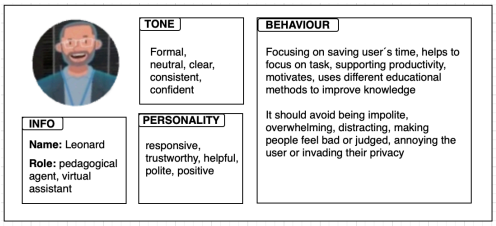

Figure 1: Agent persona. (Adapted from Figma, 2022).

AGENT FUNCTIONALITIES

For displaying system-user interactions in a system, UML use case diagrams were made using an acknowledged notation (OMG, 2005).

More than 90 functional requirements were identified based on the survey data. They were classified based on what is possible for a rule-based pedagogical agent-based avatar to do and what can be implemented through having a conversational agent. The data analysis revealed duplicates which were consolidated. Most important functionalities were determined based on the number of mentions. Considering the project limitations being able to build a rule-based agent in the form of an avatar, its limits on interacting with the user are acknowledged. As a rule-based agent, it provides support during the training scenario by saying scripted lines and includes minimal pre-programmed reactions in response to user actions. The functionalities are shown below (see Figure 2).

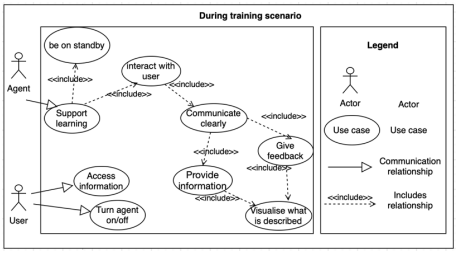

Figure 2: Use case diagram for agent functionalities.

VISUAL DESIGN PROTOTYPE

The initial version for the agent-based avatar was designed by the VR development partner within the Horizon Europe, ATHENA project (see Figure 3). The VR development partner of the project uses an agile methodology for guiding the development process. Personas and use cases developed based on survey data are a well-established approach in the industry. These outputs aim to showcase a research-based model of the agent, considering the users´ needs. The avatar represents a typical employee in the waterway and drinking and wastewater sector wearing a blue outfit

for credibility and speaking with a clear, but friendly voice. A static image resembling a hologram was chosen as a futuristic solution that contained minimal interaction for a simplified approach, such as standing in two basic positions with a happy or sad expression.

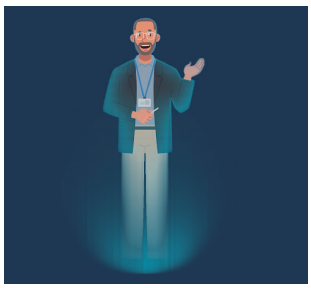

Figure 3: Avatar-based agent. Source: Virtual rangers.

DISCUSSION

In answering RQ1 on identifying important qualities for an avatar-based agent, the responders indicate the need for creating an avatar with human-like qualities, portraying an ideal teacher or a facilitator – knowledgeable and supportive. This shows that human-like interactions are desired by the target group. The target group highlights mainly the need for the agent to provide assistance to the user. It should support the learning process by being visible, interacting with the user through providing information and giving feedback about performance.

To answer RQ2 on the underlying reasons, the participants prefer to have the final say in terms of control. For example, turning the agent on or off, or configuring certain aspects related to the agent-based avatar. Also, the target group highlights the need for the agent to be beneficial for the user.

In a cybersecurity training context, the agent is seen providing the most value in a supporting role, such as giving feedback about performance and guiding the trainee during their shift.

Although this paper focuses on the functionalities of a rule-based agent in the form of an avatar, the data analysis demonstrates that the target group favors an agent that helps facilitate more interaction, such as assisting with specific tasks for the user (making summaries or a study plan etc.). Many of these functionalities can be implemented through a conversational agent that could then consider user's abilities and more.

Next to detailed responses, the survey also contained non-responses and unclear answers from the participants where content was not possible to extract. A limitation of surveys is that it does not allow researchers to ask participants to further elaborate on an answer, nor ask clarifications to better understand the responses.

This contribution enables the avatar-based agent to be designed taking into account the real needs of the target group. The results of this study inform the development process of an operational agent. The aim is for the avatar to be integrated into the training scenarios in VR.

The future work is to assess the avatar-based agent for its suitability to its intended task. In addition, considering the needs of the target group, in the next iterations, the agent-based avatar could be complemented with a conversational agent with the help of Moodle AI.

CONCLUSION

The HCD approach has been essential in understanding the needs and aspirations of the target group to better integrate innovative technologies in cybersecurity training. The agent persona and use cases give a more accurate direction to the development on how the agent-based avatars should act and be perceived by the participants going through training in VR.

ACKNOWLEDGMENT

This work received funding from the Department of Software Science at Tallinn University of Technology. This study is conducted as part of the ATHENA project (DEP: Project 101127970).

The authors have no competing interests to declare that are relevant to the content of this article.

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

Adinolf, S., Wyeth, P., Brown, R., Altizer, R., 2019. Towards Designing Agent Based Virtual Reality Applications for Cybersecurity Training, in: Proceedings of the 31st Australian Conference on Human-Computer-Interaction. Presented at the Ozchi'19: 31st Australian Conference On Human-Computer-Interaction, ACM, Fremantle WA Australia, pp. 452–456. https://doi.org/10.1145/3369457.3369515

Adinolf, S., Wyeth, P., Brown, R., Simpson, L., 2020. Near and Dear: Designing Relatable VR Agents for Training Games, in: 32nd Australian Conference on Human-Computer Interaction. Presented at the OzCHI'20: 32nd Australian Conference on Human-Computer-Interaction, ACM, Sydney NSW Australia, pp. 413–425. https://doi.org/10.1145/3441000.3441007

Baylor, A., 2011. The design of motivational agents and avatars. Educ. Technol. Res. Dev. 59, 291–300. https://doi.org/10.1007/s11423-011-9196-3

Bente, G., Schmälzle, R., Jahn, N.T., Schaaf, A., 2023. Measuring the effects of co-location on emotion perception in shared virtual environments: An ecological perspective. Front. Virtual Real. 4, 1032510. https://doi.org/10.3389/frvir.2023.1032510

Caio Calado, 2022. Chatbot or Voice Persona (Conversational Design Template) [WWW Document]. Figma. URL https://www.figma.com/community/file/1126608139029167758/chatbot-or-voice-persona-conversational-design-template (accessed 2.5.26).

Capel, T., Brereton, M., 2023. What is Human-Centered about Human-Centered AI? A Map of the Research Landscape, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI'23. Association for Computing Machinery, New York, NY, USA, pp. 1–23. https://doi.org/10.1145/3544548.3580959

ECSO WG5, 2020. Understanding Cyber Ranges: From Hype to Reality.

ECSO WG5, 2017. ECSO Gaps in European Cyber Education and Professional Training.

ENISA, 2024. 2024 Report on the State of Cybersecurity in the Union. Publications Office, LU. https://doi.org/10.2824/0401593

ENISA, 2021. Addressing the EU cybersecurity skills shortage and gap through higher education. Publications Office, LU. https://doi.org/10.2824/033355

Incident reporting [WWW Document], 2023. CIRAS. URL https://ciras.enisa.europa.eu/ciras-consolidated-reporting (accessed 2.5.26).

ISO - International Organization for Standardization, 2019. ISO 9241-210. Ergonomics of human-system interaction - Part 210: Human-centered design for interactive systems.

Martha, A.S.D., Santoso, H., 2019. The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. J. Educ. Online 16. https://doi.org/10.9743/jeo.2019.16.1.8

OMG, 2005. UML 2.0 Superstructure Specification [WWW Document]. ResearchGate. URL https://www.researchgate.net/publication/200034552_UML_20_Superstructure_Specification (accessed 2.5.26).

Taylor, T.L., 2000. The Social Design of Virtual Worlds: Constructing the User and Community Through Code.

Terry, G., Hayfield, N., Clarke, V., Braun, V., 2017. Thematic analysis. SAGE Handb. Qual. Res. Psychol. 2, 25.

Veletsianos, George, Miller, Charles, Doering, Aaron, 2009. EnALI: A Research and Design Framework for Virtual Characters and Pedagogical Agents. ResearchGate. https://doi.org/10.2190/EC.41.2.c