

Item	Response	Condition	Metric
I think that I would like to use this system frequently.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I found the system unnecessarily complex.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I thought the system was easy to use.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I think that I would need the support of a technical person to be able to use this system.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I found the various functions in this system were well integrated.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I thought there was too much inconsistency in this system.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I would imagine that most people would learn to use this system very quickly.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I found the system very cumbersome to use.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I felt very confident using the system.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
I needed to learn a lot of things before I could get going with this system.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Usability [1]
In the computer generated world I had a sense of "being there".	1. Strongly disagree 7. Strongly agree	GUI/Agent	Presence [2]
Somehow I felt that the virtual world surrounded me.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Presence [2]
I had a sense of acting in the virtual space, rather than operating something from outside.	1. Strongly disagree 7. Strongly agree	GUI/Agent	Presence [2]
How would you rate the overall realism of the VR experience?	1. Not very realistic 7. Highly realistic	GUI/Agent	Realism
How would you rate the frequency of errors encountered during the design?	1. Very rare 7. Very frequent	GUI/Agent	Errors
How severe were the consequences of the errors experienced?	1. Negligible 7. Severe	GUI/Agent	Errors
How would you describe the virtual agent's understanding of nuances and context?	Text Response	Agent	-
How would you describe the coherence and relevance of the virtual agent's responses?	Text Response	Agent	-
How consistent do you find the virtual agent's responses?	Text Response	Agent	-
Did you encounter any tasks or information the agent couldn't handle that you wished it could?	Text Response	Agent	-
How effective do you believe the system would be in a real use-case of architectural design?	1. Not effective 7. Very effective	GUI/Agent	Effectiveness
Share your thoughts on the overall experience.	Text Response	-	-

Table 1: Overview of the user study questionnaire. The first column lists the statements or questions posed. The second column presents the possible responses, which include either closed values on a 1 to 7 Likert scale or open-ended text fields. The third column specifies the condition after which the items are posed. The fourth column indicates the corresponding metrics used for evaluation, including references to relevant literature.

## References

- [1] John Brooke. “SUS: A quick and dirty usability scale”. In: *Usability Eval. Ind.* 189 (Nov. 1995).
- [2] Mehmet Berkman and Guven Catak. “I-group Presence Questionnaire: Psychometrically Revised English Version”. In: *Mugla Journal of Science and Technology* 7 (Oct. 2021), pp. 1–10. DOI: [10.22531/muglajsci.882271](https://doi.org/10.22531/muglajsci.882271).