

Evaluation Keys	NUM	Questions	Shorthand	Response Type
Presence	1	I felt engaged during the simulation.	Engagement	Likert scale
	2	I felt immersed in the computer-generated world.	Immersion	Likert scale
	3	I was able to concentrate on the simulation without being distracted by my surroundings.	Concentration	Likert scale
	4	I forgot about the real world during the interaction.	World Forgetfulness	Likert scale
Usability	5	The equipment was comfortable to use.	Equipment Comfort	Likert scale
	6	I felt comfortable interacting with the child Avatar.	Avatar Interaction Comfort	Likert scale
	7	The interface of the tool was easy to understand and use.	Interface Usability	Likert scale
	8	I did not experience technical difficulties while interacting with the child Avatar.	Technical Difficulty	Likert scale
	9	I would feel very comfortable using this tool on my own next time.	Ease of Future Use	Likert scale
Visual Fidelity	10	The appearance of the child Avatar was realistic.	Appearance Fidelity	Likert scale
	11	The virtual environment where the child Avatar was located felt real and contributed to my overall immersive experience.	Environment Fidelity	Likert scale
	12	I perceived hand-movements/gestures from the child Avatar.	Hand Movement Perception	Likert scale
	13	The quality of the child Avatar's movements was satisfactory (naturalness, realism, ...).	Movement Quality	Likert scale
	14	The child Avatar's lip movements were well synchronized with the speech.	Lip Sync Accuracy	Likert scale
	15	The child Avatar's face expressions/movements felt realistic and were well synchronized with the speech.	Facial Expression Fidelity	Likert scale
	16	The overall perception was realistic and pleasant.	Overall Realism Perception	Likert scale
Emotion	17	I felt emotionally engaged during the interaction with the child Avatar.	Emotional Engagement	Likert scale
	18	I perceived emotions in the child Avatar's responses.	Emotional Response Perception	Likert scale
	19	The child Avatar's emotional reactions (e.g. body language, facial expressions and behaviour) looked realistic.	Emotional Reaction Realism	Likert scale
	20	The child Avatar's emotional reactions (e.g. body language, facial expressions and behaviour) consistently matched the content of the interview.	Emotion-Content Match	Likert scale
Empathy	21	Interacting with the child Avatar evoked my empathy towards their situation.	Empathy Elicitation	Likert scale
	22	Please provide one or more examples of the aspects of the child Avatar that felt particularly effective in eliciting your empathy and understanding.	Effective Empathy Elicitation Examples	Open-ended
	23	Please provide one or more examples of the aspects of the child Avatar that felt particularly ineffective in eliciting your empathy and understanding.	Ineffective Empathy Elicitation Examples	Open-ended
Responsiveness	24	The responsiveness of the system to my inputs felt right, natural and smooth (e.g. the system's reaction time, the consequent responses/actions from the child Avatar).	System Responsiveness	Likert scale
	25	I noticed a delay between my questions and the child Avatar's responses/reactions.	Response Delay Notice	Likert scale
	26	The pace was the usual for a conversation with a child in such circumstances.	Conversation Pace Normalcy	Likert scale
Appropriateness	27	The child Avatar's responses felt age appropriate.	Age-Appropriate Response	Likert scale
	28	The child Avatar's responses were consistent with respect to the general story.	Story Consistency	Likert scale
	29	The child Avatar's responses were appropriate and on-topic with my questions.	Response Relevance	Likert scale
Training Effectiveness	30	From a learning perspective, my interaction with the child Avatar felt as effective as interacting with a human actor/trainer.	Training Comparability	Likert scale
	31	I think this tool should be included in investigative interviewing training programs.	Tool Inclusion Recommendation	Likert scale