

Online Resource 1: Questionnaires

Wichers, L., Schreiber, L., Möller, S., & Spang, R. P. Gaze in the Machine: How Mechanical and AI-Based Eye Contact Correction Shape Communication Experience and Gaze Behaviour in Video Calls. *Quality and User Experience*.

Corresponding author: Robert P. Spang, Chair of Usability, Faculty of Media, Bauhaus-Universität Weimar, Weimar, Germany; robert.spang@uni-weimar.de

Items were presented in German, and in English for participants who did not speak German; the English wording is reported here. Unless stated otherwise, items were answered on continuous sliders from 0 (*does not apply*) to 100 (*fully applies*). NMSPI: Networked Minds Social Presence Inventory (Biocca and Harms, 2003); UES: User Engagement Scale (O'Brien et al., 2018). The person characteristics in Sections 1 and 2 were collected to describe the sample and are not analysed in the article.

1 Pre-questionnaire (before the first condition)

1.1 Item: Weekly Video Telephony Usage

Survey Questions	Response Scale
How many hours per week do you spend on video calls and video conferences? (Private)	In 0.5h steps
How many hours per week do you spend on video calls and video conferences? (Business)	

Evaluation The answers to both questions are added up per participant to calculate the total hours of weekly video telephony usage.

1.2 Item: Replacement Potential

Survey Questions	Response Scale
Video calls can replace face-to-face conversations in the future.	Continuous slider (0 = does not apply, 100 = fully applies).

Evaluation The value of Replacement Potential (Pre) is subtracted from Replacement Potential (Post) to determine whether approval has increased (positive sign) or decreased (negative sign).

2 Post-questionnaire (after the last condition)

2.1 Item: Sympathy

Survey Questions	Response Scale
I would like to meet my conversation partner again.	Continuous slider (0 = does not apply, 100 = fully applies) for each question.
I think my conversation partner could be a friend of mine.	
I think my conversation partner is a pleasant person.	
I think my conversation partner would fit into my circle of friends.	

Evaluation The mean value was calculated for all answers to the 4 questions.

2.2 Item: Age

Survey Questions	Response Scale
How old are you?	Number in years

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

2.3 Item: Gender Identification

Survey Questions	Response Scale
How do you identify?	Multiple choice with categories: male, female, intersex/diverse, prefer not to say

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

2.4 Item: Native Language

Survey Questions	Response Scale
What is your native language?	Free-text response

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

2.5 Item: Education

Survey Questions	Response Scale
What is your highest level of education completed?	Multiple choice with categories: still in school, university entrance qualification (A-levels or equivalent), secondary school diploma (lower level), university or university of applied sciences degree, intermediate secondary school diploma or equivalent

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

2.6 Item: Profession

Survey Questions	Response Scale
What is your current main profession?	Free-text response

Evaluation Verifies whether open-ended response contains the word 'student', regardless of upper or lower case. Returns True if the word is present, False otherwise.

2.7 Item: Visual Impairment

Survey Questions	Response Scale
Do you have impaired vision?	Multiple choice question with categories: 'Yes' / 'No'
If yes: Are you currently using a visual aid?	Multiple choice question with categories: yes, single-vision glasses/contact lenses; yes, progressive glasses/contact lenses; other (see next question); no
If you use another visual aid, which one?	Free-text response

Evaluation Based on the information provided in the questionnaires, participants were divided into four categories: 'no' (no visual aid and no need), "yes" (use of a visual aid when needed), 'impaired vision' (no use of a visual aid despite existing need) and 'null' (missing or incomplete information).

2.8 Item: Technology Affinity

Survey Questions	Response Scale
Please indicate the degree to which you agree with the following statements.	Six-point Likert scale without a neutral middle category (1 = 'not at all true' to 6 = 'completely true').
I like to explore technical systems in depth.	
I like to try out the functions of new technical systems.	
I mainly engage with technical systems because I have to.	
When I encounter a new technical system, I explore it thoroughly.	
I really enjoy spending time getting to know a new technical system.	
It's enough for me that a technical system works; I don't care how or why.	
I try to understand exactly how a technical system works.	
It's enough for me to know the basic functions of a technical system.	
I try to fully exploit the possibilities of a technical system.	

Evaluation Questions 3, 6 and 8 were inverted according to Franke et al. (2019) and then the mean value was calculated across all 9 questions.

2.9 Item: Physiological State

Survey Questions	Response Scale
Are you currently hungry?	Continuous slider (0 = does not apply, 100 = fully applies) for question one and three. For question two, four and five participants could specify the number in increments of 0.5 hours.
How many hours ago did you last eat something?	
Are you currently tired?	
How many hours of sleep do you usually need to feel fit, awake, and well-rested?	
How many hours did you sleep today?	

Evaluation The results for all questions were added up.

2.10 Item: Shyness

Survey Questions	Response Scale
I feel shy in the presence of others.	Five-point Likert scale (1 = 'does not apply' to 5 = 'fully applies')
I feel inhibited around others.	
I approach other people in a relaxed manner.	
I find it easy to engage with strangers.	
I feel uncomfortable at parties and in other larger groups.	

Evaluation Question 3 and 4 were inverted according to Asendorpf (1987) and then the mean value was calculated across all 5 questions.

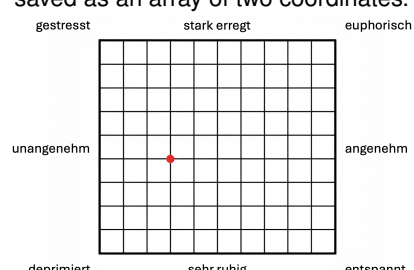
2.11 Item: Sociability

Survey Questions	Response Scale
I enjoy having many people around me.	The items were answered on a five-point Likert scale (1 = 'does not apply' to 5 = 'fully applies')
I really enjoy talking with other people.	
I usually prefer to do things on my own.	
I find people more stimulating than anything else.	
I prefer working with others rather than alone.	

Evaluation Question 3 was inverted according to Asendorpf and Wilpers (1998) and then the mean value was calculated across all 5 questions.

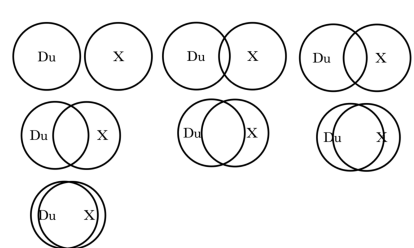
3 Condition questionnaire (after each of the four conditions)

3.1 Item: Valence and Arousal

Survey Questions	Response Scale
Please indicate how you feel at this very moment. Mark the option that best describes how you feel right now.	<p>Both axes range from 0 to 1, the mark a participant places on the graphic is therefore saved as an array of two coordinates.</p> 

Evaluation The click position is stored as proportions (0–1) of the grid width and height, measured from the top-left corner. Valence is the horizontal position (right = pleasant = high value). Arousal is one minus the vertical position, so that higher values indicate higher arousal (top of the grid = "stark erregt"). The grid was displayed with German labels: gestresst (stressed), stark erregt (highly aroused), euphorisch (euphoric), unangenehm (unpleasant), angenehm (pleasant), deprimiert (depressed), sehr ruhig (very calm), entspannt (relaxed). The red dot shows an example response.

3.2 Item: Feeling of Closeness

Survey Questions	Response Scale
Please select the picture that best describes your relationship with your conversation partner.	 <p>Inclusion of Other in the Self scale (Aron et al., 1992). Labels: Du = you; X = the conversation partner.</p> <p>Each visualization is assigned a number from 0 to 6.</p>

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.3 Item: Naturalness of Eye Contact

Survey Questions	Response Scale
The eye contact in this conversation felt natural.	Continuous slider (0 = does not apply, 100 = fully applies)
I felt at ease maintaining eye contact during the conversation.	

Evaluation Results for this category represent the average score of both survey questions.

3.4 Items: Eye Contact Intensity and Eye Contact Intrusiveness

Survey Questions	Response Scale
The eye contact in this conversation felt intense.	Continuous slider (0 = does not apply, 100 = fully applies)
The eye contact in this conversation felt intrusive.	

Evaluation The two items were barely correlated ($r = .10$) and were therefore analysed separately as Eye Contact Intensity and Eye Contact Intrusiveness.

3.5 Item: Eye Contact Anxiety

Survey Questions	Response Scale
I felt anxious because of the eye contact during our conversation.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.6 Item: Own Attention

Survey Questions	Response Scale
I felt that I remained attentive throughout our conversation.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.7 Item: Partner's Attentiveness

Survey Questions	Response Scale
I perceived my conversation partner to be attentive to me.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.8 Item: Friendliness

Survey Questions	Response Scale
I perceived my conversation partner to be warm and friendly.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.9 Item: Competence

Survey Questions	Response Scale
I perceived my conversation partner to be competent and confident.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.10 Item: Direct Eye Contact

Survey Questions	Response Scale
My conversation partner maintained direct eye contact throughout our conversation.	Continuous slider (0 = does not apply, 100 = fully applies)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.11 Item: Co Presence

Survey Questions	Category	Response Scale
I often felt as if my partner and I were in the same room together.	Self	Continuous slider (0 = does not apply, 100 = fully applies)
I think my partner often felt as if we were in the same room together.	Others	
I was often aware of my partner in the room.	Self	
My partner was often aware of me in the room.	Others	
I hardly noticed my partner in the room.	Self	
My partner didn't notice me in the room.	Others	
I often felt as if we were in different places rather than together in the same room.	Self	
I think my partner often felt as if we were in different places rather than together in the same room.	Others	

Evaluation The responses to questions 5 through 8 were reversed. Subsequently, a mean score was calculated for the category "Co Presence" across all questions. In addition, for all items in the "Self" category, the mean score for "Co Presence Self" was computed, and for all items in the "Others" category, the mean score for "Co Presence Others" was calculated.

3.12 Item: Engagement

Survey Questions	Category	Response Scale
I paid close attention to my partner.	Self	Continuous slider (0 = does not apply, 100 = fully applies)
My partner paid close attention to me.	Others	
I was easily distracted from my partner when other things were going on.	Self	
My partner was easily distracted from me when other things were going on.	Others	
I tended to ignore my partner.	Self	
My partner tended to ignore me.	Others	

Evaluation The responses to questions 3 through 6 were reversed. Subsequently, a mean score was calculated for the category “Engagement” across all questions. In addition, for all items in the “Self” category, the mean score for “Engagement Self” was computed, and for all items in the “Others” category, the mean score for “Engagement Others” was calculated.

3.13 Item: Emotional Contagion

Survey Questions	Category	Response Scale
I was sometimes influenced by my partner’s moods.	Self	Continuous slider (0 = does not apply, 100 = fully applies)
My partner was sometimes influenced by my moods.	Others	
When I was happy, my partner tended to be happy.	Others	
When my partner was happy, I tended to be happy.	Self	
When I was feeling sad, my partner also seemed to be down.	Others	
When my partner was feeling sad, I tended to be sad.	Self	
When I was feeling nervous, my partner also seemed to be nervous.	Others	
When my partner was nervous, I tended to be nervous.	Self	

Evaluation The mean score was calculated for the category “Emotional Contagion” across all questions. In addition, for all items in the “Self” category, the mean score for “Emotional Contagion Self” was computed, and for all items in the “Others” category, the mean score for “Emotional Contagion Others” was calculated.

3.14 Item: Comprehension

Survey Questions	Category	Response Scale
I was able to communicate my intentions clearly to my partner.	Self	Continuous slider (0 = does not apply, 100 = fully applies)
My partner was able to communicate their intentions clearly to me.	Others	
My thoughts were clear to my partner.	Self	
My partner’s thoughts were clear to me.	Others	
I was able to understand what my partner meant.	Self	
My partner was able to understand what I meant.	Others	

Evaluation The mean score was calculated for the category “Comprehension” across all questions. In addition, for all items in the “Self” category, the mean score for “Comprehension Self” was computed, and for all items in the “Others” category, the mean score for “Comprehension Others” was calculated.

3.15 Item: Interdependence

Survey Questions	Category	Response Scale
My actions were often dependent on my partner's actions.	Self	Continuous slider (0 = does not apply, 100 = fully applies)
My partner's actions were often dependent on my actions.	Others	
My behavior was often in direct response to my partner's behavior.	Self	
The behavior of my partner was often in direct response to my behavior.	Others	
What I did often affected what my partner did.	Self	
What my partner did often affected what I did.	Others	

Evaluation The mean score was calculated for the category “Interdependence” across all questions. In addition, for all items in the “Self” category, the mean score for “Interdependence Self” was computed, and for all items in the “Others” category, the mean score for “Interdependence Others” was calculated.

3.16 Item: Focused Attention

Survey Questions	Response Scale
I lost myself in this conversation.	Continuous slider (0 = does not apply, 100 = fully applies)
The time I spent talking to my partner just slipped away.	
I was absorbed in this experience.	

Evaluation The final score for this category represents the mean value derived from all three items.

3.17 Item: Perceived Usability

Survey Questions	Response Scale
I felt frustrated while talking to my partner.	Continuous slider (0 = does not apply, 100 = fully applies)
I found the experience of this conversation confusing to use.	
Talking to my partner was taxing.	

Evaluation The final score for this category was determined by first inverting the data for all three items and subsequently calculating their mean value.

3.18 Item: Webcam Awareness

Survey Questions	Response Scale
I consciously noticed the webcam.	Six-point Likert scale without a neutral middle category (0 = ‘does not apply’ to 5 = ‘fully applies’)

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.19 Item: Feeling Observed

Survey Questions	Response Scale
I felt observed by the webcam	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.20 Item: Discomfort due to recording

Survey Questions	Response Scale
I felt uncomfortable because I was being recorded by the webcam.	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.21 Item: Camera Gaze Discomfort

Survey Questions	Response Scale
I feel uncomfortable when looking into the webcam.	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.22 Item: Intentional Camera Gaze

Survey Questions	Response Scale
I consciously tried to look into the webcam regularly.	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.23 Item: Camera Gaze Avoidance

Survey Questions	Response Scale
I avoided looking directly into the webcam.	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.

3.24 Item: Perceived Social Closeness

Survey Questions	Response Scale
The conversation felt personal.	Six-point Likert scale without a neutral middle category (0 = 'does not apply' to 5 = 'fully applies').

Evaluation The collected data were used in their original form for the final analysis and were not further computed or transformed.