Balancing Usability and Security:

User Experience Design for Control Units in Secure Video Conferencing Systems

Isabella Åkertjärn

Supervisor: Alva Mårdsjö
Examiner: Erik Marsja

LINKÖPING UNIVERSITY

Linköpings Universitet
SE-58183 Linköping
013-281000, www.liu.se

Abstract

The knowledge of how technology is designed and evaluated is essential for creating systems that support users effectively and meaningfully, especially systems that handle high security. This study aims to explore the balance between usability and security, especially within secure video conferencing systems. The system evaluated is *Secure Conference* by Sectra Communications AB, which currently exists in a prototype format but is intended for future market release. This product is totally secure against eavesdropping and tampering, enabling safe meetings between different physical locations. Specifically, the study investigates whether users experience usability and/or workload issues with the control unit of the prototype -the part of the system that allows the user to manage the meeting, what is visually and audibly received and transmitted. The study also investigates how the relationship between high security requirements and user experience is perceived. To examine this, the study adopted a multi-method approach combining user interviews and usability testing. The interview data were analyzed using thematic analysis, and two questionnaires were used during the user tests: User Experience Questionnaire (UEQ) and NASA Task Load Index (NASA-TLX). The results indicate that usability is perceived as high and workload as low; therefore, no major issues were identified during the study. Yet, the users did provide valuable insights that could further improve the control unit, which culminated into a design proposal involving minor changes of the design.

Introduction

The global security environment is currently unstable. The Swedish Security Service (2025) emphasizes that *"The security situation remains serious. Foreign powers act in an unpredictable and deniable manner which makes the threats difficult to assess."*. In such an environment, secure communication becomes essential to prevent risks and to ensure that sensitive information cannot be exposed. However, Di Nocera et al. (2023) states in systematic literature research that usability and security often are in tension where increased security measures tend to reduce usability, and vice versa. This tension can be a huge danger - if a secure system is too difficult to use, users may choose an insecure communication method above a secure method. Ensuring both high security and high usability is therefore critical for systems that handle classified information, but it needs to be in balance.

This study will be design oriented and utilizes both quantitative and qualitative methods. It is in collaboration with Sectra Communications AB with focus on their new product: *Sectra Secure Conferencing*. The purpose of this product is to enable the exchange of classified information through secure video communication, without the risk of eavesdropping or tampering from unknown actors. Thus, it is essentially a highly protected version of a conventional video conferencing system. The Secure Conference is designed to operate between two physically separated locations and to handle classified information up to the security level SECRET (Sectra Communications, n.d.). This is the second highest classification level out of four (Swedish Defence Materiel Administration, 2022). The system will be ensured since it builds on the high-performance encryption Sectra NetEncryptor, which according to Sectra Communications (n.d.) is allowed to handle *"Real-time transfer of TOP SECRET data"*. The product is intended to primarily be used by governmental agencies, defense organizations, and companies that handle sensitive information.

The system is designed in a relatively simple way: there will be two rooms (Figure 1) set up to have a meeting. Each room is equipped with three screens displaying video feeds and shared computer screens. A speaker and a camera are placed in each room, as well as a control unit (Figure 2) that allows users to receive and share visual content and manage the audio transmission. When the control unit is powered on via the top-right button, it illuminates green (Figure 3) and the flip-switches illuminate red when activated (Figure 4). This control unit is the specific component that this study evaluates.

Figure 1

The content of the conference setup

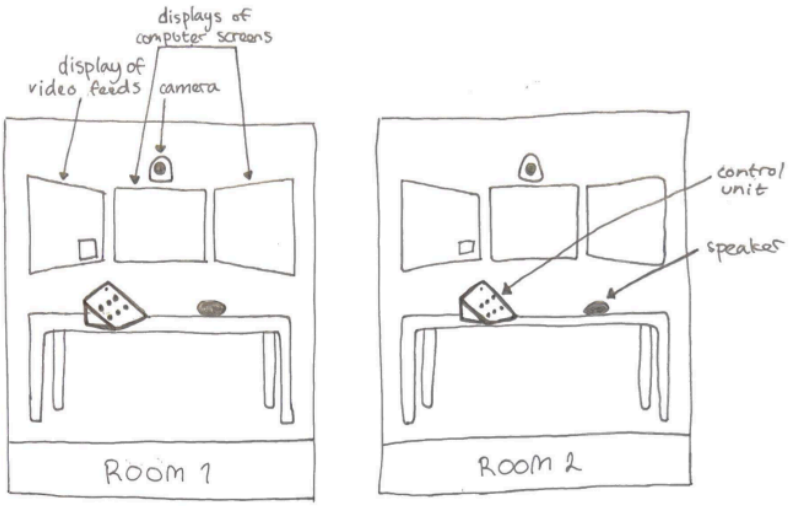


Figure 2

The current prototype of the control unit



Figure 3

Power button when powered on



Figure 4

Flip-switch when activate



No special technical expertise is needed from the end user, as the meetings will be configured by an "administrator" in advance allowing the user to enter the room and attend the meeting. This means that the administrators have the key role for the technical aspects of the conference system, such as maintenance and troubleshooting. The administrator is also required for security reasons, since it is security-classified information that will be handled during the meeting. However, even though the user does not need to set up the meeting it is still essential that the system remains intuitive for the user *during* the meeting.

Furthermore, the secure conference systems must operate in special rooms that prevent eavesdropping and signals to leek. The Swedish Defence Materiel Administration (2009) explains that TEMPEST-rooms are enclosed by a metallic shield that blocks unintentional electromagnetic signals that occur when technological equipment processes, stores or transmits information. Signals can otherwise be analyzed or intercepted by unauthorized organizations and reveal sensitive information.

There are of course advantages to having meetings in person, for example a more natural flow of the conversation, the help to reduce misunderstandings and no technical problems. On the other hand, remote meetings also provide benefits, including reduced travel time, lower costs, and fewer logistical challenges. Since we do know that governmental, military, and other

organizations that need to talk about security-classified subjects often are placed in different physical locations. Furthermore, the GREENER principles (Lannelongue et al., 2023) promote environmentally responsible choices with the goal of reducing the climate footprint. Where remote meetings can be understood as a preferable alternative due to their lower climate impact compared to long-distance travel. Therefore, the opportunity to have remote meetings is very important no matter how secret the topic may be.

In addition, *workload* is a factor that may increase the risk of users choosing less secure systems over secure systems. Hart and Staveland (1988) describe workload as a subjective, multidimensional entity that reflects the perceived cost incurred by an operator to achieve a particular level of performance. Workload is not considered an inherent property of a task itself, since it emerges from the interaction between task demands. The conditions under which the task is performed, and the individual's capabilities, strategies, and perceptions. It is therefore a human-centered construct, capturing how a person *experiences* the demands of a task rather than simply what the task objectively requires.

Research Statement

The aim of this study is to investigate how the design of the control unit can be optimized to ensure high usability and low workload, critical aspects for a security classified product. This includes what usability problems may exist in the current prototype, how workload may be reduced, as well as how the high security requirements might affect usability. Specifically, this study will examine this through interviews and user tests with the goal of producing a design proposal of the control unit. The research questions are as follows:

RQ1: What usability problems do users experience with the current prototype of the control unit?

RQ2: How does the control unit prototype affect the participants perceived workload?

RQ3: How do participants perceive the relationship between high security classification requirements and user experience?

Methods

The methodology of this study consisted of semi-structured interviews, followed by user tests. Since it combines qualitative and quantitative data the study employs a mixed method approach. According to Gilbert & Stoneman (2015), mixed methods research involves two or more methods in one study and is advantageous as it gives the researchers a chance to understand their topic better, in various directions. This also increases the accuracy and the level of confidence in research.

Participants

The same six participants (hereafter referred as P1-P6) took part in both interviews and user tests, in addition to one participant who was involved in a pilot test. The participants had diverse professional backgrounds (two engineers, two thesis workers, a research manager and a contract manager) and were all employees at Sectra. However, not all of the participants had previous knowledge about the system, which was considered beneficial since the system later will be used by people not only with pre-knowledge. The Company uses Microsoft Teams as its primary video conferencing system in daily work.

The recruitment was conducted via email, which informed about the purpose of the study, that the participation was voluntary, and that they could withdraw from the interview at any time without consequences. This information is according to the Swedish research council (Swedish Research Council, 2002) four main principles regarding research ethics:

- *Information Requirement:* All participants were informed about the purpose and conditions of the study to make an informed decision about participation.
- *The Consent Requirement:* Participants gave verbal consent initially.
- *The Confidentiality Requirement:* The participants privacy was protected by anonymizing all responses, and by storing it with access restricted to the researcher.
- *The Requirement of Use:* All collected data were used solely for the purpose of this study, and the data not used in this report will be deleted upon the study's completion.

Materials

For the *interviews* the only material needed was the interview guide (Appendix A). During the *user tests* the prototype of the conference system was set up between two different buildings in Mjärdevi. This prototype included all necessary parts for conducting the tests, such as screens, speakers, camera and the control unit. The users were provided with different *scenarios* (Appendix B) and completed two surveys: *User Experience Questionnaire* (Appendix C) and *NASA Task Load Index* (Appendix D).

User Experience Questionnaire (UEQ)

User Experience Questionnaire (n.d.) offers a survey designed to evaluate the experienced usability of products. Since it is a quite long survey consisting of 26 statements this method

provides a comprehensive overview covering multiple aspects of the perceived usability. The statements are divided into six categories:

1. *Attractiveness:* The overall impression of the product.
2. *Perspicuity:* The intuitive experience of the interface.
3. *Efficiency:* How effective the product feels, regarding to complete tasks.
4. *Dependability:* If the product feels secure and predictable.
5. *Stimulation:* If the user experiences the product as exciting and motivating to use.
6. *Novelty:* The products appear as creative and innovative.

Schrepp (2023) explains that each statement uses a semantic differential format, for example "attractive - unattractive", on a 7-point scale. Saad et al. (under review) show in their methodology-guideline that UEQ is established and comprehensive as it scored 85% out of 100%. This score is judged by experts to be suitable for assessing user experience which indicates trustworthiness.

NASA Task Load Index (NASA-TLX)

The NASA-TLX measures subjective workload especially in systems that regard human-computer interactions. The method is multi-dimensional since it is built on six subscales, each rated between 0-100, where 0 is very low and 100 is very high workload (except for Performance that has a reversed scale) (Hart, 1986). A meta-analysis by Grier (2025) shows that the scores vary between different studies and work fields, but that many systems have a typical score of approximately 35-55. Hart (1986) presents the subscales as following:

1. *Mental Demand:* How mentally tough the assignment was.
2. *Physical Demand:* How physically tough the assignment was.
3. *Temporal Demand:* How time-pressured the pace of the task was.
4. *Performance:* How successful the task was.
5. *Effort:* How hard the participant needed to work to accomplish the task.
6. *Frustration:* How irritated and stressed did the task make the participant feel.

Cao et al. (2009) claims that NASA-TLX has been tested and validated across a variety of fields, for example laboratory experiments, supervisory tasks and flight simulators. This indicates that the questionnaire is established and widely used in systems that measure mental workload. However, this study is specifically using *Raw NASA-TLX*, that Said et al. (2020) explains as participants completing all the ratings in the questionnaire, but without weighing them. In the original NASA-TLX the participants need to weigh the subscales towards each other, but research indicates that the raw version correlates strongly with the weighted version, while also being more time-efficient and easier to administer.

Procedure

This section describes the procedure of the study, including the semi-structured interviews and user tests. The pilot participant completed the exact same procedure as outlined below and is therefore not described in detail.

Interviews

The study began with *semi-structured interviews* (Appendix A), aimed to explore the participants' prior experiences and expectations regarding the conference system. Semi-structured interviews mainly follow the interview guide but allow for flexibility and follow-up questions (Merriam & Tisdell, 2016). These were subsequently analyzed with *thematic analysis*, which Braun och Clarke (2020) reports as a method to identify, analyze and report recurring patterns and themes with the goal to discover meaningful inferences. It involves six systematic phases to help the researcher: First, the researcher should familiarize themselves with the collected data by reading and, if necessary, transcribing it. Second, initial codes are generated to identify meaningful segments. Third, the researcher searches in the codes to organize main themes and sub-themes. Fourth, themes are reviewed for relevance, coherence, and representation of the data. Fifth, themes are defined and named with detailed descriptions. Finally, the analysis is reported by relating the themes to the research questions.

User Tests

Following the interviews, the same participants took part in user tests to evaluate the current prototype. Participants were asked to complete four scenarios (detailed in Appendix B):

1. Start the conferencing system and verify that a secure meeting has been established. The purpose is to assess how efficiently the user can confirm that the meeting is connected.

2. Transmit and receive audio and video correctly. The purpose is to evaluate the participants' experience to turn audio and video on and off, as well as understanding whether they are shared and received.

3. *C*onnect a computer to the system, transmit and receive screen sharing via the control unit. The purpose is to test how easily the participant can start screen sharing, understand what is being shared and receive.

4. The participant needs to end the meeting quickly. The purpose is to determine if the participant can terminate the meeting immediately (may be necessary in an emergency), and to confirm that no audio, video, or screen content is being shared or received after shutting down.

During the scenarios the participants were encouraged to *think-aloud*. According to McDonald et al. (2012) this method can identify usability problems by verbalizing thoughts,

impressions, and actions in real time. After completing the scenarios, participants were asked a few post-test questions (Appendix B) to capture additional reflections and finalize their feedback. At the end of each user test session, participants completed UEQ that measures perceived usability and NASA-TLX that measures experienced workload.

The UEQ were analyzed in an Excel-sheet where the researcher entered the questionnaire scores. The 7-point scale was coded from -3 to +3, with -3 being the most negative, 0 neutral and +3 the most positive (Schrepp, 2023). The NASA-TLX scores were calculated using the online NASA-TLX Calculator (n.d.). The researcher entered each participant's responses on the website, after which an excel-file was generated containing all answers and the raw NASA-TLX scores ranging from 0-100. Based on the calculated scores, Hart & Staveland (1988) explain that the researcher can interpret results regarding what subsets had higher or lower effects on cognitive load.

Results

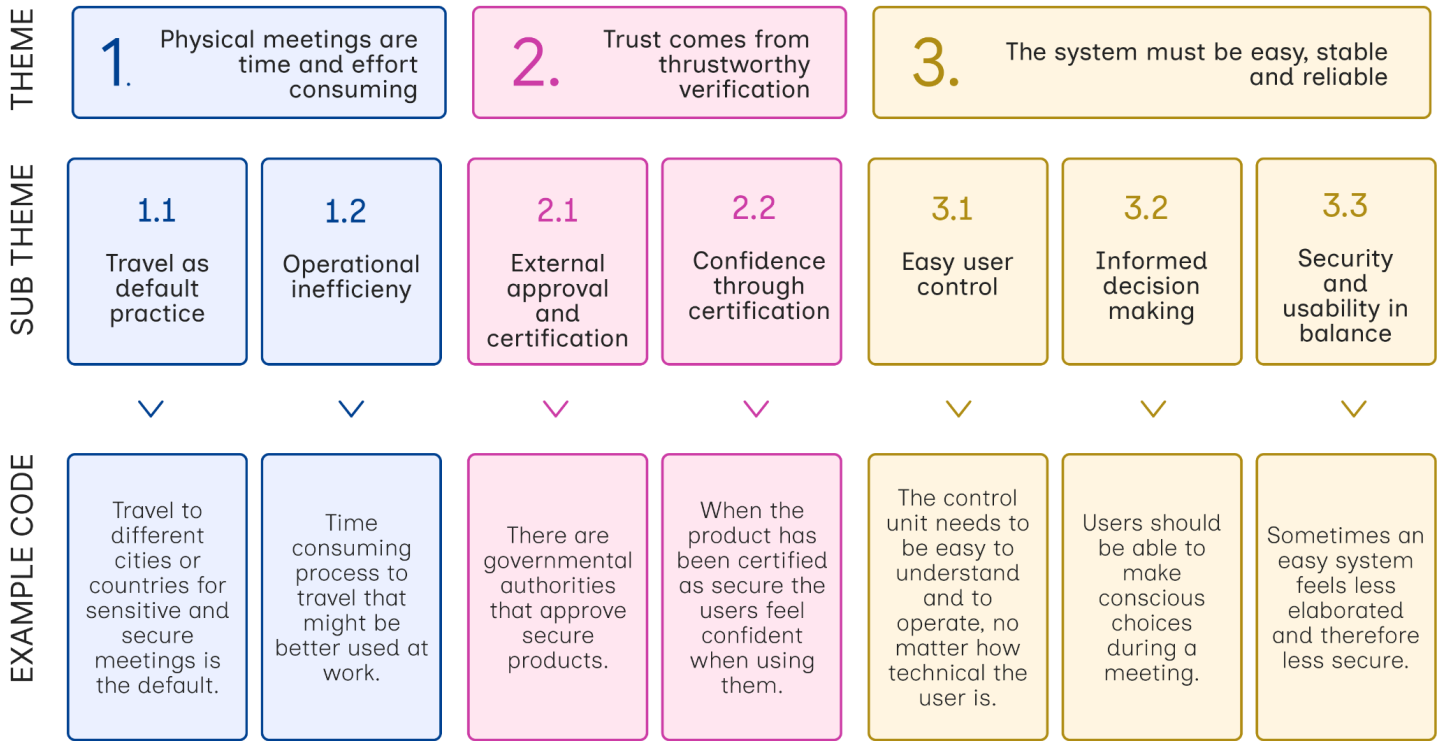
This chapter presents the results from the interviews and the user tests, including a thematic analysis and results from UEQ and NASA-TLX questionnaires. Also, additional findings regarding design changes are presented in a structured list.

Interviews

The semi-structured interviews were analyzed using an thematic analysis method, resulting in three main themes with associated subthemes and example codes (Figure 5). The themes are furtherly described and illustrated with representative quotes from participants below.

Figure 5

Themes, sub themes and example codes from the thematic analysis



1. Physical meetings are time and effort consuming

This theme captures the participants' descriptions of how secure meetings are often conducted in physical locations, requiring travel to other cities or even countries. Two subthemes further illustrate this theme. First, *Travel as default practice, r*eflects that traveling to attend meetings is the standard approach, regardless of whether the meeting itself is short or could potentially be handled remotely. Second, *Operational inefficiency,* highlights how this can reduce overall efficiency. Workdays that are consumed by travelling, participants rather spend doing productive work in the office. The quote *"We often need a full nine-hour workday to go to Stockholm to have a two-hour meeting." (P3)* explains how a relatively short meeting may require an entire workday due to travel time, illustrating the mismatch between effort invested

and meeting duration. Participants' experiences suggest that physical meetings may hinder productivity and highlight the value of an alternative meeting method.

2. **Trust comes from trustworthy verification**

The second theme reflects participants expressed thoughts about trust in secure systems, together with two subthemes: *External approval and certification* and *Confidence through certification.* The first subtheme highlights that external governmental authorities validate and certify secure products before they are approved for use. The second subtheme emphasizes how the external certification generates confidence, providing reassurance that the system is secure and that potential risks have been mitigated. The quote *"I trust that our products are totally secure only after they are approved and verified." (P1)* suggests that confidence in a system stems from external approval, before approval participants described uncertainty regarding potential inference or tampering.

3. **The system must be easy, stable and reliable**

An intuitive, stable, and reliable system is essential for ensuring that users can engage with the system without frustration or errors. Three subthemes further elaborate this theme. First, *Easy user control*, highlights that the system must be easy to operate regardless of the users technical capabilities. Second, *Informed decision making,* reflects the importance of users making conscious choices during meetings. Third, *Security and usability in balance*, participants expressed that easy systems sometimes appear less elaborated and therefore also less secure. The quote *"It must be easy to use. We need non-technical people to want to use it, to avoid them to rather use less secure options." (P5)* illustrates that a too complicated system may impact users to rather use insecure alternatives. At the same time, participants did acknowledge that simplicity in usability must not compromise with the security of the system.

User Tests

Results from the UEQ are shown in Table 1 and Figure 6. Attractiveness (M=1.44) was classified as *Above average* according to the UEQ benchmark and rated as 25% of better results and 50% of worse results. Perspicuity (M=2.58), Efficiency (M=2.33) and Dependability (M=2.58) were rated as *Excellent* placing them among the top 10% of results. Stimulation (M=0.54) performed *Below Average* with 50% of better results and 25% of worse results. Novelty (M=-0.75) received the classification *Bad*, placing it among the range of the 25% worst benchmark results.

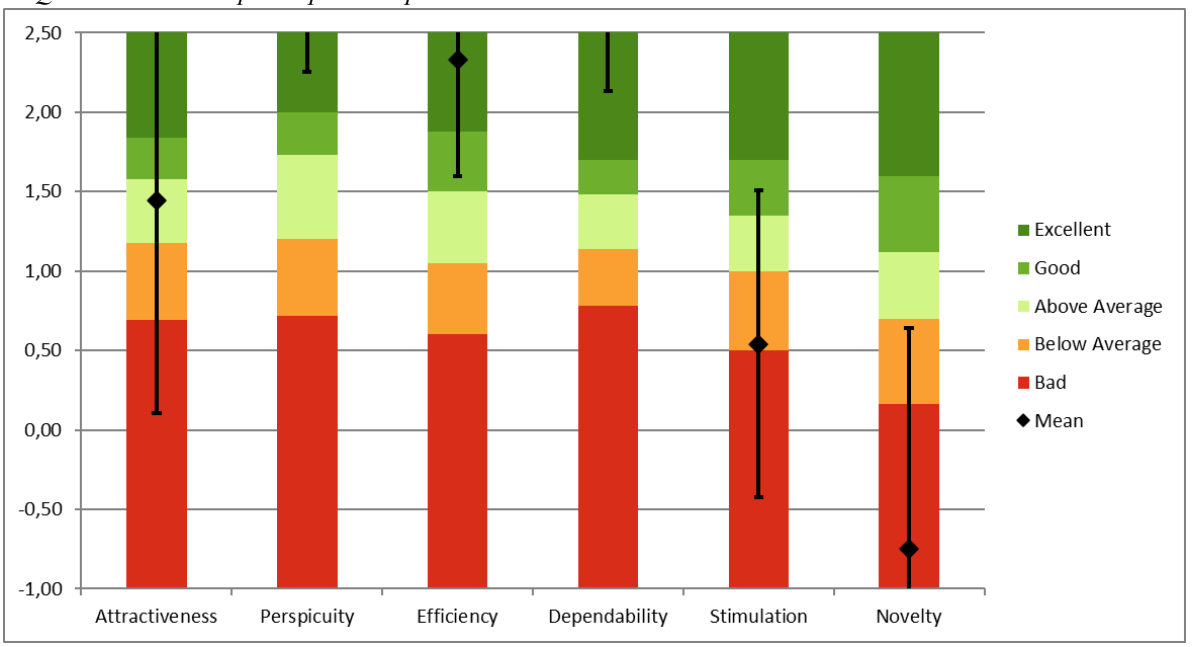
Table 1

UEQ results based on participants responses

Scale	*Mean*	*Comparison to benchmark*	*Interpretation*
Attractiveness	1.44	Above average	25% of results better, 50% of results worse
Perspicuity	2.58	Excellent	In the range of the 10% best results
Efficiency	2.33	Excellent	In the range of the 10% best results
Dependability	2.58	Excellent	In the range of the 10% best results
Stimulation	0.54	Below Average	50% of results better, 25% of results worse
Novelty	-0.75	Bad	In the range of the 25% worst results

Figure 6

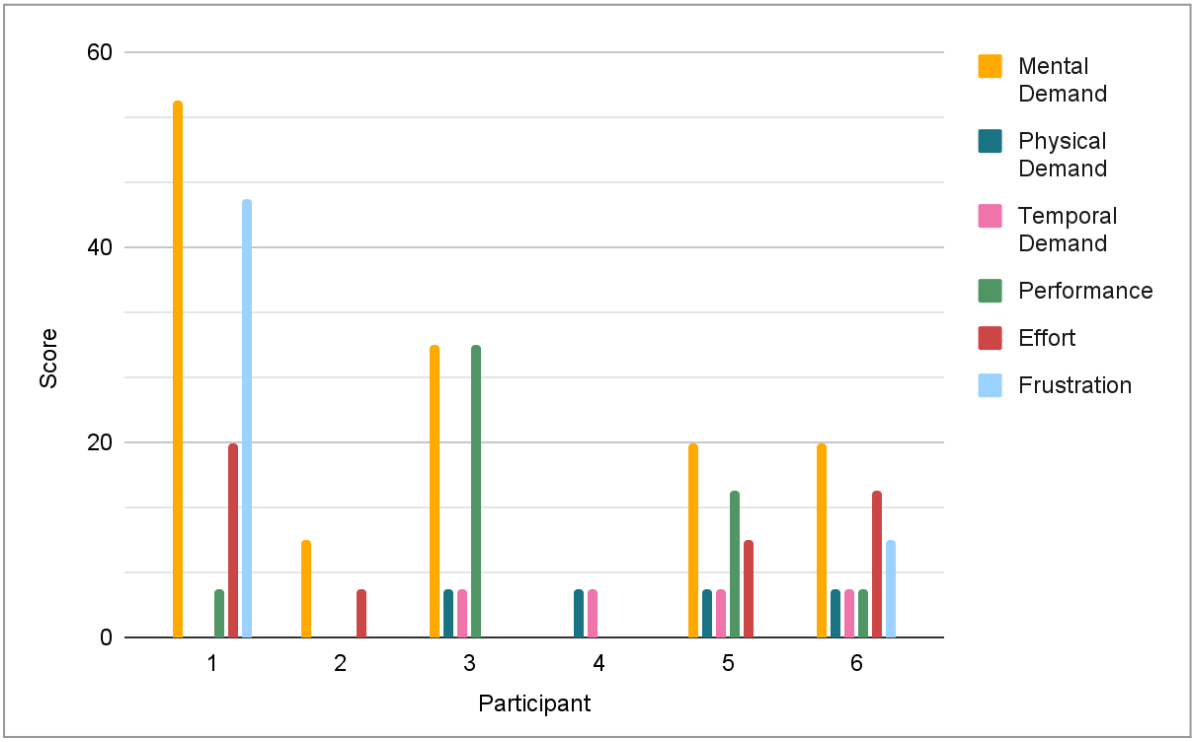
UEQ results based on participants responses



NASA-TLX results are presented in Figure 7. Generally, it shows low answers, with 13 out of 36 answers not shown in the diagram since they had a score of 0. However, P1 reported higher scores on Mental Demand (55) and Frustration (45). P3 reported higher scores on Mental Demand (30) and Performance (30). The remaining ratings had a score of 20 or lower.

Figure 7

Participants responses regarding NASA-TLX categories



Since this study used the unweighted NASA-TLX (Raw TLX) the workload score for each participant was calculated as the mean of the six subscales, with the following formula:

$$Raw\ TLX = (MD+PD+TD+P+E+F) / 6$$

The individual Raw TLX scores are presented in Table 2. P4 and P2 reported lowest scores (1.67 and 2.50). P5 and P6 reported moderate scores (9.17), while P3 showed slightly higher scores (11.67). The highest reported Raw TLX score was P1(20.83).

Table 2

The participants' Raw Score for the study

Participant	*Raw TLX Score*
P1	20.83
P2	2.50
P3	11.67
P4	1.67
P5	9.17
P6	9.17

The *Mean Raw TLX Score* of 9.17 was calculated by averaging the individual scores, using the formula:

$$Mean\ Raw\ TLX = \sum Participant\ Raw\ TLX\ Score\ /\ N\ participants$$

Additional Findings

During the user tests the participants were instructed to think-aloud and subsequently they answered post-test questions, this established the following additional findings:

Power button:

- The button was too small; a larger finger can make it hard to press down.
- When the button is pressed down the system instantly shuts on/off, which is clearly indicated when the screens and the light on the control unit turns off.

Flip-switches:

- The flip-switches were appreciated regarding the tactile feel in them and were described to clearly lock in position, reducing the risk of accidental switching.
- Add icons to clarify the function for each switch.
- Reorganize the switches according to typical usage order: (1) Camera, (2) Audio, (3) PC Screen.
- Illuminate green when activated and illuminate red when deactivated. This can show a bigger difference between on/off.

Additional observations:

- The overall design of the control unit was appreciated and one participant even expressed that it felt easier to use than a conventional conferencing system (e.g. Microsoft Teams or Zoom) because it has fewer options.

Discussion

In the following chapter, the results of the study are discussed with the aim of answering the research questions. In addition, the study's methodology is analyzed.

Result Discussion

To address the first research question, *RQ1*, the results from the UEQ are analyzed with the list of additional findings. *RQ2* is addressed through the NASA-TLX results and *RQ3* is addressed through the themes identified from the thematic analysis.

Research Question 1

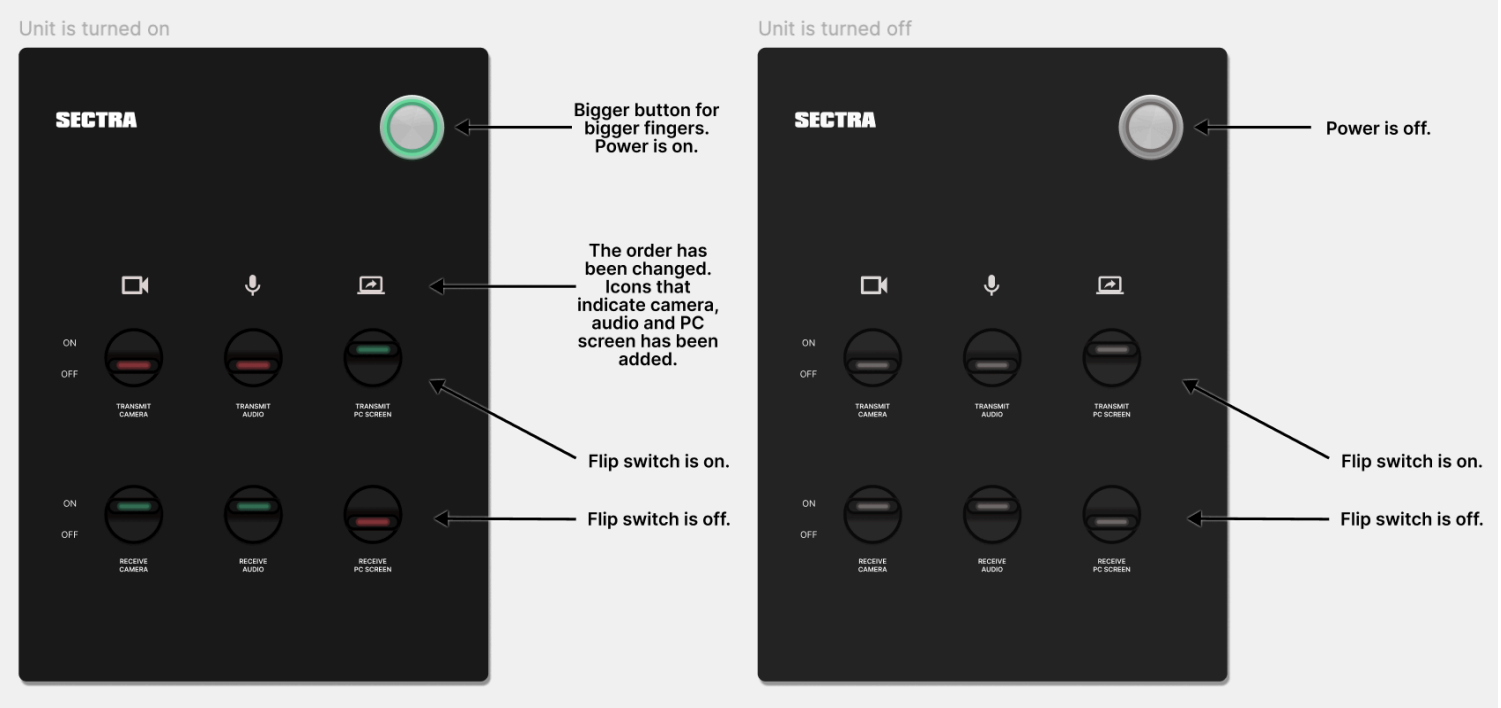
The first research question - *What usability problems do users experience with the current prototype of the control unit?* - is addressed as follows: A product of this nature must be easy to use so that users genuinely want to use it. The UEQ results indicate that there are no major usability issues in the control unit prototype. The category *attractiveness* scored *above average,* reflecting a positive impression of the interface. *Perspicuity*, *efficiency* and *dependability* were all rated *excellent*, suggesting users found the product intuitive, tasks were completed quickly, and the system felt secure and predictable - aligning with the products intended purpose. The categories *stimulation* and *novelty* performed *below average* and *bad*, respectively. However, this was expected and not necessarily negative. A secure, safety-critical product should not be perceived as innovative or exciting, rather as predictable and trustworthy. These results indicate that the current design functions well regarding usability.

The additional findings suggest small improvements could further enhance clarity and intuitiveness, leading to the design proposal in Figure 8. The mockup presents two versions of the control unit - powered on (left) and powered off (right). Key changes are marked out with comments, including:

1. An enlarged, green-lit power button when on, unlit when off.
2. Added icons above each flip-switch clarifying their functions.
3. Rearranged switches into a more logical sequence based on typical usage.
4. Green illumination when switches are activated, red when deactivated, and no light when powered off.

Figure 8

Mockup of the control unit with small changes where the left one is turned on, and the right one is turned off



One usability problem may occur if a flip-switch remains in the “ON” position after a meeting. Since switches only have two physical positions, a switch can remain physically up even when the system is powered off. This could be problematic if, for example, audio transmits unintentionally at the start of a new meeting. To prevent this, administrators should turn off all switches after each session to ensure every meeting begins with no function activated. This is an operational consideration requiring appropriate routines rather than an usability flaw.

The qualitative data identified improvements such as adding icons to flip-switches and changing illumination to green when activated. Physical, tactile interaction with buttons and switches supports user confidence through direct feedback. It is important to avoid excessive visual stimulation by limiting colors and avoiding overly prominent icons. Operational routines must also be considered. Figure 8 strengthens visual communication through icons and color indicators but should be tested and evaluated before further development.

Research Question 2

The second research question - *How does the control unit prototype affect the participants perceived workload?* - is addressed through the NASA-TLX results. Overall scores were very low, indicating participants experience no notable workload difficulties. Individual scores ranged from 1.67 to 20.83, with a mean Raw TLX score of 9.17. On a 0-100 scale, these results reflect a generally low perceived workload. According to Grier (2025), typical global NASA-TLX scores range between 35 and 55, meaning the average score of 9.17 in this study is comparatively low. This suggests the current prototype does not impose substantial

cognitive or physical demands on users and does not significantly affect participants' perceived workload.

Figure 7 shows that 13 of 36 responses scored 0. P1 experienced higher mental demand (55) and frustration (45), while P3 experienced higher mental demand and performance scores (30), indicating the system might feel demanding and frustrating to these specific participants. Although these scores deviate from the otherwise low ratings, they should be interpreted cautiously. A plausible explanation is that the prototype, being in an early development stage, occasionally required users to interpret interface elements not yet fully refined. Despite this, the overall pattern remains clear; most participants experienced minimal workload across all six NASA-TLX dimensions, indicating the design aligns with the goal of enabling secure communication without unnecessary workload.

An important consideration is whether further measures could reduce workload even more. The study's scenarios involved a simulated video meeting where the system functioned without issues, but under such conditions users are less likely to experience workload or stress. In contrast, if the system were used during an actual meeting with technical difficulties, users would likely experience higher stress levels, potentially leading to higher reported workload in the NASA-TLX results. This aspect should be considered in future studies.

Research Question 3

The third research question - *How do participants perceive the relationship between high security-classification requirements and user experience?* - is addressed through the thematic analysis and the additional findings.

The theme *"Trust comes from trustworthy verification"* indicates that participants feel confident a system is fully secure only when it is approved and certified by external authorities. While the system's design is expected to communicate seriousness and professionalism, perceived security is primarily grounded in formal verification rather than interface design alone. This aligns in previous research by Di Nocera et al. (2023), which highlights that usability and security often are in tension where increased security measures often reduce usability, and vice versa. This tension introduces potential risks, as overly complex systems may lead to users choosing an easier but less secure method instead. Therefore, achieving a balance between security and usability is essential, particularly in systems handling classified information.

Furthermore, the theme *"The system must be easy, stable and reliable"* emphasizes that usability remains essential regardless of the user's technical expertise. This highlights the importance of intuitive interaction and operational clarity, facilitating quick decision-making. For example, the system should allow users to easily turn off the system if an unauthorized individual enters the room. At the same time, an overly simplified interface may risk being perceived as less secure because they appear less complex. Participants expressed a need for

the system to feel both easy to use and robust, indicating that the design must communicate stability, reliability, and security.

Additional findings show that the control unit was generally perceived as intuitive, with one participant meaning that it was easier to use than systems like Microsoft Teams or Zoom due to its reduced options and buttons. The UEQ results further reinforce this perception, showing excellent ratings in perspicuity, efficiency, and dependability while stimulation and novelty scored lower. Similarly, the NASA-TLX results indicate a low perceived workload (mean score 9.17) suggesting the system felt easy to operate despite its security requirements.

In summary, the relationship between high security-classification and user experience requirements is complex, requiring a balance that makes users feel both secure and willing to use the system. Participants expect the system to be intuitive and reliable without appearing overly simplistic and trust that security will be ensured through certification of the product. Therefore, it is essential to consider design in the development process to ensure usability is balanced with the high security demands that inevitably influence the final product.

Methods Discussion

In this section, the methodological choices throughout the study are presented and discussed as regards to participants, materials and the procedure.

Participants

The study included six participants and one additional pilot test user. This sample size is considered acceptable and sufficient for usability testing. According to Jakob Nielsen (2000) the sample size of six participants is sufficient to identify the majority of usability issues, according to. Testing with five participants can reveal approximately 85% of usability problems, as each additional participant tends to uncover fewer new issues due to diminishing returns.

Furthermore, even though all participants work at Sectra, they came from different educational and professional backgrounds. This is seen as a positive aspect of the participant selection since they have different expertise and can provide different points of views. However, it would have been valuable to test the system on actual end users, such as military personnel or politicians. Their perspectives would have provided more authentic insights, but accessing such users is of course highly difficult and were not feasible within the scope of this study. Still, not all the participants had prior knowledge of the system, and they performed tasks similar to real-world usage. They could therefore provide meaningful feedback on the system's usability and functionality, since many future users will probably also lack prior experience. However, a limitation may be that all participants do use Microsoft Teams in their daily work and actual end users may not use Teams or any digital conferencing systems at all. This difference may influence how the system is perceived; for example, features that appear basic to experienced users may be perceived differently by inexperienced users.

It is considered positive that the same participants took part in both the interviews and the user tests, since this allowed their thoughts from the interviews to be explored further while interacting with the prototype. This approach was also more time effective regarding, for example, recruitment. The participants were recruited using the same process and provided informed consent, according to the four research ethics principles established by the Swedish Research Council (2002). Following these ethical guidelines ensures the research was conducted in accordance with established research ethics standards.

Materials

The prototype used during the user tests are the real hardware prototype currently under development, which has also been used in demonstrations with external companies. This is considered as a strength of the study, in contrast to using a low- or high-fidelity prototype, since it increases the ecological validity of the study. For example, participants commented positively on the tactile feel of the flip switches and noted that the power button was too small for comfortable use. Such insights would likely not have been shared if a non-physical or simplified prototype had been used.

To evaluate the participants' user experience the User Experience Questionnaire (n.d.) was used, which provides a comprehensive overview of several usability aspects. Additionally, Saad et al. (under review) report that UEQ scores highly (85%) regarding methodological establishness and comprehensiveness. This supports the suitability of the method for this study. The NASA-TLX was used to measure participants' subjective workload. It has been validated across a wide range of domains, indicating it is a well-established method for measuring mental workload (Cao et al., 2009). The subscale of performance is reversed compared to the other subscales, placing “good” where the others indicate “low” (Appendix D). This requires participants to read the questions carefully and may reduce the risk of answering quickly or inattentively. However, this study used the *raw* version of NASA-TLX. Said et al. (2020) explain this as a simplified form of the original method. In the full version, participants weigh each subscale based on personal opinion of importance, but the raw NASA-TLX results correlates strongly. In this study, the full version could have provided additional interesting insights, but the choice to use the raw method is considered reasonable within the scope of the study, since it does provide correlating results.

Procedure

Semi-structured interviews were conducted to gather qualitative insights from the participants. According to Merriam and Tisdell (2016), this method is flexible and allows the researcher to ask follow-up questions and adapt the discussion to each participant's experiences and perspectives. This enabled the interviews to generate varied insights rather than repetitive responses. The data were analyzed using thematic analysis following the approach by Braun och Clarke (2020). This analysis method generated concrete and meaningful themes to this study. However, since thematic analysis relies on the researcher’s interpretation, there is a risk of subjective bias. Different researchers analyzing the same

material may identify different themes and the reliability could therefore be questioned. It would have been beneficial if multiple researchers had independently coded the data and compared the results, but this was not possible due to the course requirements.

The user tests were based on scenarios representing the four most critical functions of the system, according to the researcher. While additional scenarios could have been included, they would likely have yielded limited additional insights, given that the product itself offers relatively few functions. In addition, a limitation of the test setup may be that only one side of the conference interaction was directly used, since the participant and researcher were located in the same room. Observing both sides would likely been valuable to gather experiences when having an actual conversation from both ends of the system, but it would also require an additional researcher. Also, incorporating more stressful scenarios, such as situations where technical issues occur (e.g., the audio not functioning), would likely have provided different NASA-TLX scores and further insights into the product's functionality under stressful conditions.

During the user tests, participants were encouraged to think-aloud. As described by McDonald et al. (2012), this method can reveal usability issues by asking participants to verbalize their thoughts while performing tasks. Although the method may influence user performance (since talking while performing a task can be distracting), this was considered a minor problem in the study since the tasks were not time-based and therefore probably not experienced as stressful. Therefore, the think-aloud method most likely provided deep insights into participants' experiences that might not have been captured otherwise.

Conclusions

This study explored how the design of a control unit can support a balance between usability and security. The UEQ indicated no major usability issues, and the NASA-TLX suggested that the workload required to operate the prototype was generally low. During the user tests, participants expressed several positive impressions of the current prototype. For example, the overall layout of the control unit and the tactile feel of the flip-switches was well received. However, several minor improvement suggestions that could enhance the user experience were also revealed. These insights resulted in a design proposal of the control unit that introduces small - but potentially important - changes. Even so, this proposal should be seen only as a preliminary suggestion and presumably requires further testing before progression to development. Moreover, a key aspect of the study was understanding how usability can be balanced with strict security requirements. The thematic analysis showed that users' trust mostly stems from reliable verification and certification by governmental authorities. This suggests that while the interface should communicate seriousness and professionalism, the feeling of security is primarily based on official certifications rather than design alone. Finally, the study's findings indicate that the product has strong potential, but slight changes might further improve the design. This is why it is important to consider design in the development process - to ensure usability is balanced with the security demands that inevitably influence the product.

AI Use Disclosure

In this study, AI (ChatGPT) was used to correct grammatical and sentence-structure errors. Since this was the first big assignment I ever wrote in English, AI was very helpful to ensure I had a consistent and accurate use of language throughout the text.

References

Braun, V., & Clarke, V. (2020). *One size fits all? What counts as quality practice in (reflexive) thematic analysis?* Qualitative Research in Psychology, 18(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Cao, A., Chintamani, K., Pandya, A. K., & Ellis, R. D. (2009). *NASA TLX: Software for assessing subjective mental workload*. Behavior Research Methods, 41(1). https://doi.org/10.3758/BRM.41.1.113

Di Nocera, F., Tempestini, G., & Orsini, M. (2023). *Usable Security: A Systematic Literature Review*. Information, 14(12). https://doi.org/10.3390/info14120641

Gilbert, G.N. & Stoneman, P. (Eds.) (2015). *Researching social life*. (4th ed.) Los Angeles: SAGE.

Grier, R. A. (2015). *How high is high? A meta-analysis of NASA-TLX global workload scores*. ResearchGate. https://www.researchgate.net/How-high-is-high-A-metanalysis.pdf

Hart, S. G. (1986). *NASA task load index (TLX): Computerized version (Volume 1.0)*. NASA Research Center. https://ntrs.nasa.gov/api/citations/20000021487/downloads/20000021487.pdf

Hart, S. G., & Staveland, L. E. (1988). *Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research*. Elsevier BV. https://doi.org/10.1016/s0166-4115(08)62386-9

Lannelongue, L., Aronson, H. E. G. & Bateman, A. (2023) *GREENER principles for environmentally sustainable computational science*. Nature Computational Science. https://doi.org/10.1038/s43588-023-00461-y

Mcdonald, S., Edwards, H. & Zhao, T. (2012). *Exploring Think-Alouds in Usability Testing: An International Survey. IEEE* Transactions on Professional Communication. https://doi.org/10.1109/TPC.2011.2182569

Merriam, S.B., & Tisdell, E.J. (2016). *Qualitative Research: A Guide to Design and Implementation*. Jossey-Bass: A WileyBrand. https://download.e-bookshelf.de/download/0003/7195/84/L-G-0003719584-0007575839.pdf

NASA-TLX Calculator. (n.d.). *NASA Task Load Index*. https://nasa-tlx-calculator.vercel.app

Nielsen, J. (2000). *Why you only need to test with 5 users*. Nielsen Norman Group. https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Saad, L., Roesler, E., Phillips, B. K., & Trafton, J. G. (under review). *Choosing the "perfect" scale: a primer to evaluate existing scales in HRI*. https://hriscaledatabase.psychology.gmu.edu/

Said, S., Gozdzik, M., Roche, T., Braun, J., Rössler, J., Kaserer, A., Spahn, D., Nöthiger, C., & Tscholl, D. (2020). *Using the raw NASA Task Load Index questionnaire to assess perceived workload in patient monitoring tasks: A meta-analysis using a mixed model*. https://doi.org/10.2196/preprints.19472

Schrepp, M. (2023). *User Experience Questionnaire Handbook (Version 11)*. UEQ-online. https://www.ueq-online.org/Material/Handbook.pdf

Sectra Communications. (n.d.). *Sectra NetEncryptor*. Retrieved January 15, 2026, from https://communications.sectra.com/products/sectra-netencryptor/

Swedish Defence Materiel Administration. (2009). *RÖS - Röjande Signaler [RÖS - Compromising Signals]*. https://www.fmv.se/globalassets/dokument/om-fmv/sakerhetsskydd/ros-rojande-signaler-2009.pdf

Swedish Defence Materiel Administration. (2022). *Industrisäkerhetsskyddsmanual [Industrial Safety Protection Manual]*. https://www.fmv.se/globalassets/dokument/om-fmv/sakerhetsskydd/ism-2022.pdf

Swedish Research Council. (2002). *Forskningsetiska principer inom humanistisk-samhällsvetenskaplig forskning [Research ethics principles in humanities and social science research]*. https://www.vr.se/download.pdf

Swedish Security Service. (2025, 22 January). *Allvarligt säkerhetsläge i en osäker tid [Serious security situation in a time of uncertainty]*. https://www.sakerhetspolisen.se/ovriga-sidor/nyheter/nyheter/2025-01-14-allvarligt-sakerhetslage-i-osaker-tid.html

User Experience Questionnaire. (n.d.). https://www.ueq-online.org

Appendix

Appendix A - Interview guide

Introduction

Hello, and thank you very much for taking the time to participate in this interview! The purpose of this session is to gain a better understanding of what you know about Sectra Secure Conference and which features you perceive as most important. Are you familiar with what the system is? *(If not, provide a brief explanation here, e.g. a secure version of Teams that uses cryptography.)*

Before we begin, I would like to ask for your consent to collect and store the data generated during this session. All collected data will be anonymized and handled confidentially. The data will be deleted at the conclusion of the study, with the exception of anonymized material that is retained and presented as part of the study's results. You can withdraw from the study at any time.

Does that sound good? Then let's get started!

Interview Questions

Background and Experience:

1. What do you know about Secure Conference?
2. How do you currently handle situations where sensitive information needs to be shared? Have you previously used any types of secure communication systems?
    a. How secure did you feel when using them?
    b. What worked well / what worked less well?

Priorities and Values:

3. What do you consider most important in a secure video communication system?
    a. Functionality? Speed? Clear security information? Low learning curve?
4. Should a secure alternative be as intuitive as a commercial solution?
5. To what extent do you think usability can be compromised for security?
    a. Are there situations where such compromises are more acceptable?
6. Which features do you consider absolutely essential in a secure video conferencing system?

Trust and Security:

7. What makes you trust a technical system? What makes you distrust one?
8. How would you like to receive confirmation that the communication is secure?
9. Can you describe a situation where you felt insecure in digital communication?

Appendix B - User Test guide

Introduction

Thank you for taking the time to participate in this user test! You will carry out a number of tasks that are typical of situations where sensitive information needs to be shared between two geographically separated locations. During the tasks, I would like you to think out loud - describe what you see, what you expect to happen, and any thoughts or questions that come up. Afterwards, I will ask you a few post-test questions to wrap this up!

Before we begin, I would like to ask for your consent to collect and store the data generated during this session. All collected data will be anonymized and handled confidentially. The data will be deleted at the conclusion of the study, with the exception of anonymized material that is retained and presented as part of the study's results. You can withdraw from the study at any time.

Does that sound good? Then let's get started!

Scenarios

Scenario 1: Initiating a Secure Meeting

- Start the video conferencing system.
- Verify that a secure meeting has been established between this room and another room.

Scenario 2: Managing Audio and Video During

- Activate your microphone.
- Confirm that audio is received from the other room.
- Activate your camera.
- Confirm that video is received from the other room.

Scenario 3: Sharing and Receiving Screen Content

- Connect a computer to the conferencing system and initiate screen sharing via the control unit.
- Receive the other rooms screen sharing.

Scenario 4: Rapidly Terminating the Meeting

- Imagine: “A person enters the conference room and you must immediately terminate the meeting”.
    - Terminate the meeting and confirm that no audio, video, or screen content is being transmitted or received.

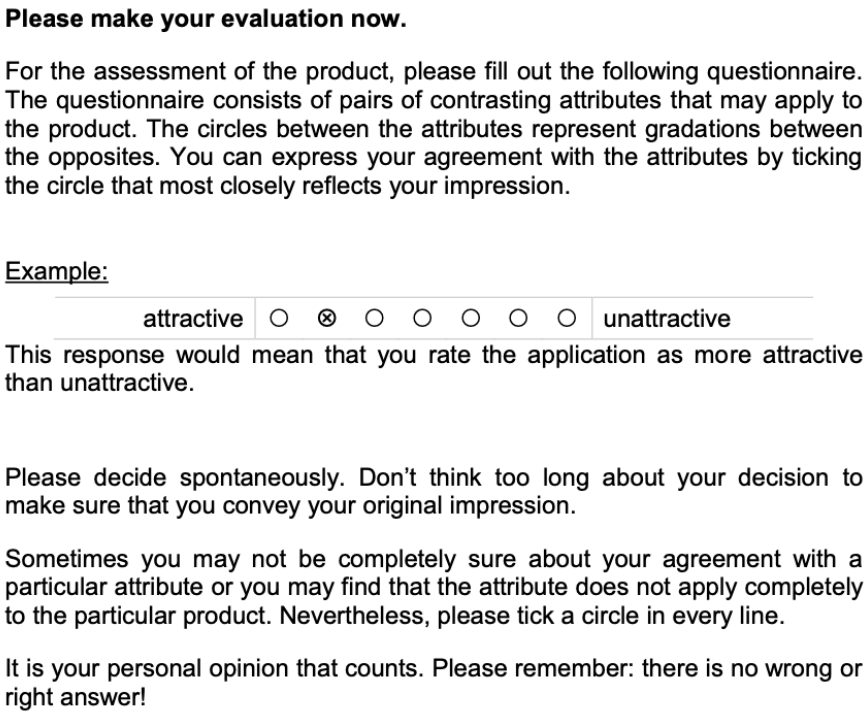
Post-Test Questions

1. How did you experience using the conferencing system?
    a. Was anything unexpectedly difficult or easy?
    b. Did the perceived level of security feel sufficient?
2. Was it easy to understand what the other room could see and hear?
3. Were there any moments during the test when you felt unsure about what was happening?
    a. What would have made it clearer?
4. Did you feel confident that the meeting was secure?
    a. What signals or indicators would be needed to strengthen that sense of confidence?
5. How does this system compare to other video conferencing systems you use?
    a. In what ways is it better or worse?
6. Would Secure Conferencing fit into your work?
    a. Or can you imagine a situation where it would be useful?

Appendix C - User Experience Questionnaire

Figure 9

User Experience Questionnaire (UEQ) retrieved from https://www.ueq-online.org



Please make your evaluation now.

For the assessment of the product, please fill out the following questionnaire. The questionnaire consists of pairs of contrasting attributes that may apply to the product. The circles between the attributes represent gradations between the opposites. You can express your agreement with the attributes by ticking the circle that most closely reflects your impression.

Example:

attractive	○	⊗	○	○	○	○	○	unattractive

This response would mean that you rate the application as more attractive than unattractive.

Please decide spontaneously. Don't think too long about your decision to make sure that you convey your original impression.

Sometimes you may not be completely sure about your agreement with a particular attribute or you may find that the attribute does not apply completely to the particular product. Nevertheless, please tick a circle in every line.

It is your personal opinion that counts. Please remember: there is no wrong or right answer!

Figure 10

User Experience Questionnaire (UEQ) retrieved from https://www.ueq-online.org

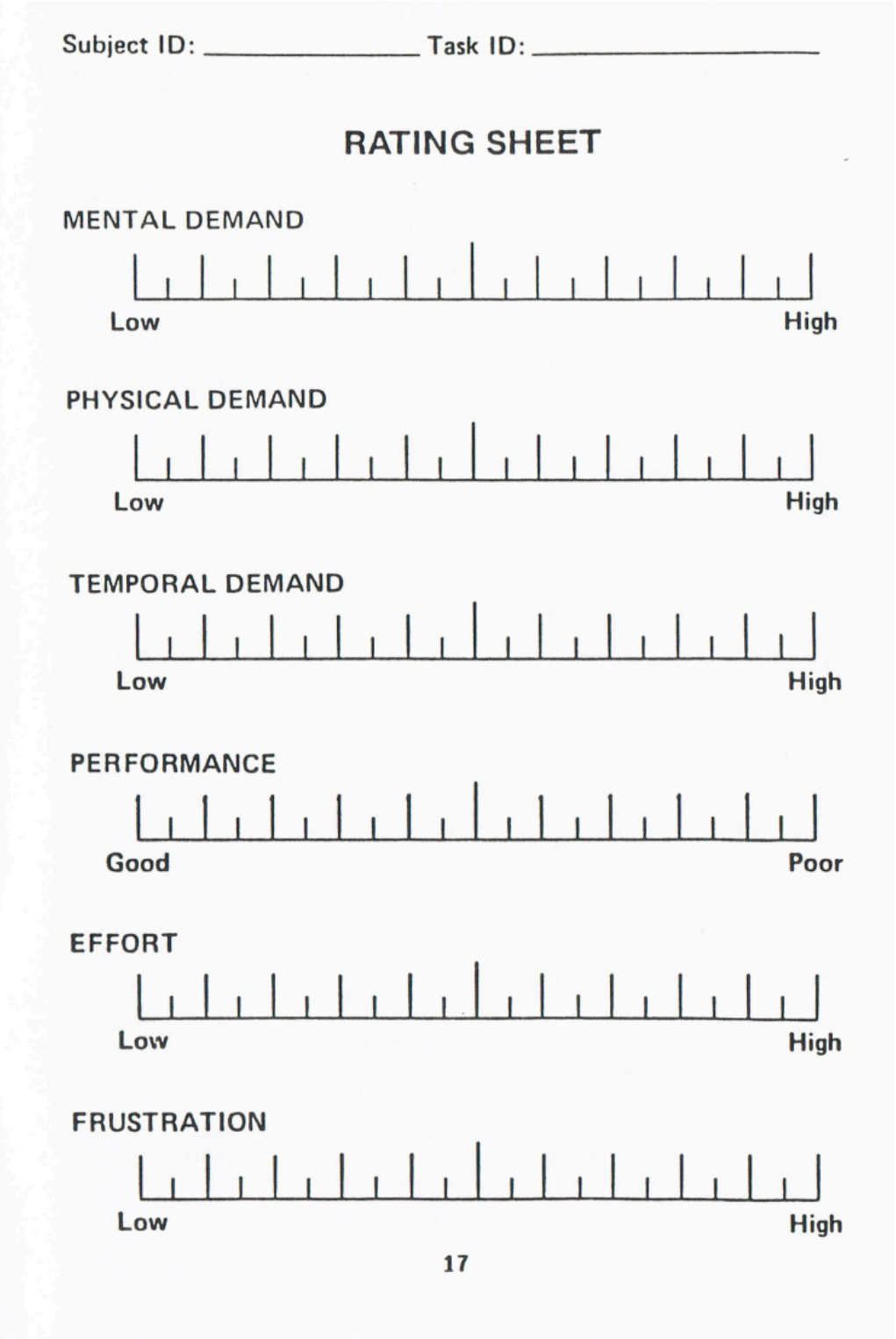
Please assess the product now by ticking one circle per line.

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Appendix D - NASA Task Load Index

Figure 11

NASA Task Load Index retrieved from TLX_pappen_manual.pdf



Appendix E - Individualised learning goals

1. *Vad ska du lära dig? (mål)*

Jag kommer att utveckla ett välgrundat förslag på hur kontrollenheten i Secure Conference kan förbättras med avseende på användarvänlighet och arbetsbelastning. Specifikt är mina mål att:

- Identifiera och analysera problem med användarvänlighet och arbetsbelastning i den befintliga prototypen.
- Förstå hur användare upplever relationen mellan höga säkerhetskrav och användarupplevelse.
- Ta fram ett designförslag som balanserar användarvänlighet och säkerhet.
- Utveckla mitt akademiska skrivande på engelska - eftersom detta första engelska rapporten jag skriver.

2. *Hur ska du lära dig det? (resurser och strategier)*

För att uppfylla mina mål kommer jag att:

- Genomföra intervjuer för att förstå användarnas behov och upplevelser.
- Utföra användartester för att identifiera konkreta problem med användbarhet och arbetsbelastning.
- Analysera insamlad kvalitativ data med tematisk analys.
- Utveckla designförslag baserat på insikter från intervjuerna och användartesterna.
- Delta i kursens peer-review-seminarier och använda feedbacken för att förbättra arbetet kontinuerligt.
- Använda engelsk litteratur för att få stöd i mitt engelska skrivande.

3. *Måldatum för färdigställande*

Senast den 28 maj ska hela rapporten vara färdig, i samband med den postersession som ska hållas då.

4. *Hur ska du veta att du lärt dig det? (bevis)*

Jag kommer att visa att jag har uppnått målen genom att:

- Producera en rapport som tydligt redogör studiens metod, resultat och analys.
- Presentera identifierade problem med användbarhet och arbetsbelastning.
- Presentera ett designförslag baserat på studiens resultat.

5. *Hur ska du bevisa att du lärt dig det? (verifiering)*

Jag kommer att verifiera att jag har uppnått målen genom att:

- Tydligt redogöra för resultaten från intervjuer och användartester.
- Genomföra en tematisk analys.
- Diskutera resultaten kritiskt.
- Presentera ett välmotiverat designförslag.

6. *Handledare feedback (utvärdering)*

- Välgrundade diskussioner kring metodval samt iterativt arbete kring deras tillämpning i studien. De tekniska förutsättningarna i produktens uppsättning gjorde testning utmanande men Isabella hittade ett protokoll som stöttade testning trots detta.
- Proaktiv vid bokning av avstämningsmöten och deras innehåll.
- Förslaget på omdesign tog avstamp både i användarnas behov som identifierades vid studien samt företagets behov och möjlighet att ändra design. Förslagen är realistiska att genomföra och skulle ge stor förbättring i upplevelsen.

Alva Mårdsjö, 26/5-26