Glorialba Rivera Marcano

Testing PreViS:

Usability Evaluation of a Prehospital Emergency Telemedicine System.

Master’s thesis in Interaction Design
Supervisor: Eleftherios Papachristou
Co-supervisor: Giovanni Pignoni
December 2022

Master’s thesis

NTNU
Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

NTNU
Norwegian University of
Science and Technology

Glorialba Rivera Marcano

Testing PreViS:

Usability Evaluation of a Prehospital Emergency Telemedicine System.

Master's thesis in Interaction Design
Supervisor: Eleftherios Papachristou
Co-supervisor: Giovanni Pignoni
December 2022

Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

NTNU
Norwegian University of
Science and Technology

Abstract

The Norwegian Health Center Sykehuset Innlandet is currently exploring the potential of integrating live video conference capability into emergency response/paramedic activities through a system called PreViS. First-hand research revealed that, even though the project is in the soft-launching stage within a growing number of paramedic stations in the region, no usability studies have been made on it. Because of this, the current research was dedicated to designing, conducting, and analyzing such studies on the device.

The main research questions were:
Q1: Can the system help make examinations more accurate in comparison to the control group?
Q2: Is the usability of the system acceptable for its use in prehospital emergencies?

The principal method used to answer Q1 was that of medical simulations. Two groups of paramedics —with 3 participants each—, conducted the same simulated emergency scenario. The control group followed the standard operating procedures expected of their practice in Norway. The intervention group used the PreViS headset instead. The data from this simulation was sifted through a performance evaluation framework created for this purpose.

To answer Q2, a usability inspection was performed on a) behaviors observed during the simulation and, b) the interfaces of the headsets (including GUI, and voice-command / navigation structure). The results of this evaluation were compared against usability heuristics.

After these methods were implemented, the results were the following:
For Q1, it was found that the simulations in the intervention group performed more accurately than those in the control group, the latter having been marred by errors and consequent harmful interventions. The visual component of the headset, in particular, helped make the examination more accurate, bypassing the interpretation of the interloper, and providing an almost first-hand point of view for the Remote Healthcare Professional (RHP).

In the case of Q2, the findings were less positive. Several heuristic violations and system errors were found on the inspection. The most notable of them was the presence of a technical failure

which made the device respond randomly to voice commands when no such action had been taken. The appearance of these was very disruptive, with 8 random activations appearing in one simulation alone. Other findings were related to a subpar heuristic implementation of the GUI, navigational elements, and system language and nomenclature.
The conclusion of this research question is that —in the opinion of the author—, the current iteration of the system is suffering from too many usability violations to be ready for launching, especially for such a complex hectic environment as the emergency prehospital practice. An overhauling of the software of the device is necessary for it to be able to fulfill usability principles.

Q1 was answered positively, while Q2 was answered negatively. These results are however not incompatible. If the results of this research are found to be generalizable when submitted to larger, more diverse sampling, it would mean that the idea of PreViS has a good proof of concept, as it would make examinations more accurate, and reduce the potential of patient harm. The flaws found on PreViS were all rooted in the imperfections of the device itself, not in the concept. Even if the issues found on the evaluation cannot be amended in future iterations, there is always potential that the idea of PreViS can find a more suitable home on a different device. The rise of wearables, AR, and body-mounted cameras make this prospect a very attainable one.

Acknowledgements

I would like to thank the following for their support on this endeavor:

To G. Pignoni and E. Papachristou for their time and good advice.

To Nils, for his kindness, inspiring dedication and patience.

To Daniella, Kevin, Eira, Annie, and Clara, for their unwavering support.

To N. Gaiman for giving me stunning universes to rest my head on.

Finally, to my family: Mi madre y mi tia. Las dos personas más fuertes que jamás he conocido, gracias por enseñarme implacabilidad, y llenarme de fortaleza.

Index.

List of Figures.

List of Tables.

List of Abbreviations.

GDPR: General Data Protection Regulation
GUI: Graphic User Interface
HCI: Human Computer Interaction.
IA: Information Architecture
IoT: Internet of Things
NIHSS: National Institutes of Health Stroke Scale.
NSD: Norsk senter for forskningsdata (Norwegian Centre for Research Data).
PHP: Pre-Hospital Professional.
PM: Paramedic.
RHP: Remote Healcare Professional
R&D: Research and Development
SUS: System Usability Scale.

1. Introduction

The current research was conducted in order to contribute to the development of a new Teleconference System dedicated to the Prehospital context (PreViS).

PreViS is a Norwegian project form Sykehuset Inlandet, which intends to use Visual Collaboration Telemedicine (i.e video conferencing tools) to connect paramedics attending the scene of an emergency with remote Healthcare professionals who can provide decision support on complex cases that need it.

Visual Collaboration Telemedicine technology such as PreVis has the potential to become a game-changer for the paramedic profession, as well as the entire emergency chain response. However, it was observed in early stages of this inquiry that the development of this tool is being made within a professional silo, and no usability or Human Computer Interaction (HCI) studies are being conducted to evaluate the performance of this system in what is already a very complex context. This is unfortunately not a rare occurrence: Research and Development (R&D) cycles of medical technologies can often be conducted in the vacuum of hermetic professional practices which, can tend to overlook the complexities of overlapping realms relevant to the end product, such as of usability, HCI, or even more intricate areas such as the complex social ecology of medical work environments. Maurin Söderholm offers an interesting perspective as to why this behavior seems to be prevalent in the design of medical tools:

“ [...] *technologies for medical contexts in particular seem to still be developed within an engineering tradition, with much focus on technical issues and the use of quantitative outcome measures as primary evaluation of systems’ effectiveness and success. This is, to some extent, understandable because technologies for health care and medical settings are developed to be operating in environments that are heavily impacted by strict requirements for safety and reliability, and rules and guidelines for patient integrity and data security; and, also because of the evidence-based approach dominating the medical field overall. However, in order to develop technologies that are truly useful, justified and successful, the practices and perspectives of users and other stakeholders also needs to be taken into account.”*

(Söderholm, M. 2013 P.6).

Söderholm's point of view highlights the motivation for participating in this research. As previously mentioned, when R&D processes are developed in technical silos, they can often miss the importance of involving Human Centered methodologies, leading to products and systems which are alien to their users and can actually hinder their performance. In the context of PreViS, this is a particularly unfortunate practice, due to the fact that within the exercise of prehospital medicine the margin of error is minimal and the consequences can be catastrophic.

1.1. Relevance for Society and the Field of Interaction Design.

Visual Collaboration Telemedicine systems like PreViS have the potential to significantly improve the effectiveness of the care given to patients, and provide 'relief' to the heavy cognitive and emotional load prehospital professionals may experience while treating complex patients. **The successful implementation of these types of systems could bring a tremendous benefit for society at large. In contrast, a defective launch could 'poison the well' and prevent further development not only for PreVis, but for possible similar Telemedicine projects as well.**

If the elements of the system end up hindering the overall performance of the prehospital professional due to poor usability elements (e.g. if the functionality of the system puts extra cognitive load on the paramedic, dividing their attention from the patient, or the audio/ video provided by the system is not clear enough to be useful for remote dialogue or diagnosis), then the odds of success will lower. However, if PreViS' usability is thoroughly observed, pain points can be identified, highlighted and properly addressed in following iterations, this will open the path for the system to achieve its expected potential, possibly creating an unprecedented positive effect on the prehospital healthcare chain, the wellbeing of their workers, and possibly life-saving impact on the patients. In other words, the success of this concept can be more achievable if there is a thorough, objective, and well documented observation of the usability and HCI dimension of the project. **Because of this, the current research is interested in evaluating the usability of the current iteration of PreViS and creating usability evaluation frameworks which can help detection of errors, in order to facilitate future analysis and strategies of mitigation.**

From a different point of view, this research can also be a stepping stone for future inquiries in interaction design for similar telemedicine applications. The early stages of this research were

inspired by the paper “The potential impact of 3d telepresence technology on task performance in emergency trauma care” by Hanna M. Söderholm, Diane H. Sonnenwald, et.al (2007), which dealt with a similar concept to the PreViS project, albeit framed in theoretical technology. Söderholm’s research design was the first inspiration for the Usability Evaluation Framework that was created for this thesis. Their work, much like this research, falls under the definitions of intermediate-level knowledge in Interaction design research, where the academic exercise is focused on creating “generative knowledge” i.e. “knowledge that plays a direct role in the creation of new designs” (K. Höök and J. Löwgren Pp. 23:2).
It is an aspiration that this study can also bequeath some useful knowledge to future researchers wishing to evaluate usability and HCI in the complex context of telemedicine.

1.2. Research Questions

This research conducted medical simulations on two groups of participants (control and intervention) in order to compare and analyze their interactions. The control group worked with the methods currently implemented by the paramedics working in Innlandet; the intervention group worked using PreVis. The observations of this exercise were then submitted through an Usability Evaluation Framework. The purpose of this was to answer the following research questions.

Q1: Can the system help make examinations better in comparison to the control group?

Q1.1. Can complex examinations be performed more accurately in comparison to the control group?

- Was the accuracy of the examination higher when done through PreVis?

Q1.2. Can the system help avoid patient harm?

- Are there fewer harmful interventions with PreVis, compared to the control group ?

Q1.3. Can the symptomatology of the patient be effectively transmitted to the RHP?

- Can all signs and symptoms that the patient is presenting be captured on video?
- Is the audio and video quality of the device high enough for the RHP to make an educated decision?

Q2: Is the usability of the system acceptable for its use in prehospital emergencies?

Q2.1. Do users perceive the system to be easy to use?

Q2.2. Is the use of the device causing an extra load (cognitive or otherwise) on the user.

- Does the setup of the device respond to the user's mental model? Are there consistent user errors?

Q2.3. Is the HCI intuitive? (voice-command interaction, navigation etc).

- Are the touchpoints of the interaction intuitive?
- Does the language used in the system (voice commands, sections labeling), match real world conventions?
- Is the system's navigation clear and accessible to the average user?

Q2.4. Is the interface prone to user errors?

- Does the interface violate usability heuristics?

Q2.5. Is the system prone to technical errors?

- Are there consistent malfunctions in the system?

As it can be seen these questions aim to evaluate and understand if the tool can actually deliver on its potential, as well as its level of usability and performance in HCI matters. The output of answering these questions will hopefully allow for an understanding of its performance and what are the areas if any that need attention for future iterations.

2. Background.

2.1.General Context: Prehospital Care and Telemedicine.

Prehospital care is defined as the medical care provided between the scene of an emergency to the transportation and handover of the patients to an appropriate healthcare facility. (Söderholm, H, 2013). Within recent decades the level of care and expected competence and sophistication of the paramedic practice has increased significantly (Summers, A. 2010). What once was expected to be basic stabilizing procedures and an emphasis on fast transportation to a health center, has advanced into pre-diagnosis and first-stage treatment before and during transportation. The tasks of transitioning between prehospital care and in-hospital care involve a web of very complex interactions weighing on the shoulders of paramedics: from patient-paramedic communication, pre-diagnosis and treatment, use of highly specialized tools and equipment, and multidisciplinary communication and coordination with remote teams. All these activities happen within critical time frames and stakes which often involve life or death outcomes. The success of this balancing act is central to the successive links to the chain of care (Söderholm, M. 2013).

As previously mentioned, this research will be based inside the early launching stage of a telemedicine project called PreViS, which would allow prehospital teams to communicate via live video-conferencing technology with specialized medical teams in a remote location; this communication bridge promises to provide medical decision support for critical patients either at the scene and/or inside the ambulance, amplifying the range, speed and quality of treatment to the patients.

After the COVID 19 pandemic, the pervasiveness of teleconference as a problem-solving tool might make this proposal seem like a deceivably simple fix. However, when it comes to telemedicine in the prehospital environment, simple fixes are few and far between.

Telemedicine has been defined as the "*the use of electronic information and communications technologies to provide and support health care when distance separates the participants*"

(Field, 1996, p.1),(Sonnenwald et al. 2008). Telemedicine is an umbrella term which covers a diverse range of technology, from health records management, SMS-based patient appointment reminder systems, and medical imaging transmission, to systems that facilitate remote surgery. In terms of Visual Collaboration Telemedicine (i.e. Video conference-calls technology in a healthcare setting), most of the current documented implementations have a focus on consultation in a non-emergency setting[1], leaving only minimal applications that are published and properly documented in the prehospital environment (Rörtgen, et al. 2013), (Skorning, M., et. al. 2012.), (Espinoza, A.V, et. al. 2016), (Pedrotti, et.al. 2021). The reasons for this might be multifocal. First, it was only until the last decade when the adoption of Third Generation devices which facilitated the wireless capacity for live streaming of video achieved sufficient resolution to be adopted for broadcasting useful medical visual data from the field. (Espinoza, A.V. et. al. 2016.). Secondly, the design considerations for emergency healthcare are exceedingly complex: *[Emergency healthcare] is characterized by time-critical, dramatic and unpredictable conditions and the involvement of different actors and organizations implies collaborative challenges and information gaps as well as loss in the overall patient care process. Furthermore, [...] working conditions and practices in different contexts such as hospitals, accident scenes or rural medical centers are sometimes unclear or badly understood among its actors and organizations. This negatively affects overall collaboration, information flow, and the coordination and use of resources.* (Söderholm, M. 2013 pp. 3).

2.2. State of the Art.

It has been difficult to pinpoint the state of the art on Visual Collaboration Telemedicine in a prehospital environment. A literature review has highlighted sparse relevant results with sufficient documentation. The majority of the publications found were early stage conceptualizations which were tested through controlled simulation studies (Rörtgen, D., Bergrath, S., et al. 2013), (Skorning, M., et. al. 2012.), (Cicero, M.X., et. al. 2016). The most complete, well documented solution found during this research was the 'PreSSUB system', a pilot study conducted in Belgium in 2016. Self-proclaimed as "The first ever 24/7 in-ambulance telemedicine service" (Espinoza, A.V., et. al. 2016.), the proposal consists of in-ambulance telemedicine, bidirectional video streaming limited to prehospital care for stroke. It was pilot

[1] This is excluding the onslaught of implementations that were brought to the frontlines due to the Covid-19 pandemic, which while very interesting, may have lack of sufficient published documentation to be relevant in this research.

tested in 2015 to seemingly positive results, albeit very little quantitative or qualitative data regarding feedback from the users/participants was shown in the study, and results shown seem to be more focused on KPIs and the technical performance of the equipment. Further research is needed to understand the current state of this project, and its performance on a User Centered/HCI perspective.

Fig. 2.01. The PreSSUB 3.0 system.
From the project description: “The teleconsultant uses a lightweight laptop computer to access the telemedicine platform. Consultations can be performed from virtually any location, as long as mobile broadband connectivity is available (a). The telemedicine device is securely mounted to the ceiling of the ambulance (b)”. (Espinoza, A.V. 2016 p.19).

2.2.1 Norwegian and Swedish Context.

2.2.1.1. The Status Quo.

First-hand interviews the researcher has conducted with prehospital professionals at Sykehuset Innlandet and prehospital professional trainers at NTNU Gjøvik conveyed that the current way paramedics are getting decision support from remote physicians is through mobile phone calls conducted either at the scene of the emergency, and/or during the ambulance transportation.
Common examples of how these calls occur is as follows.

1. The Paramedic at the scene finds a patient with an unclear or complex symptomatology.

2. The Paramedic uses a mobile phone to call a Remote Healtcare Professional (RHP) in order to get decision support (the RHPs can be General Practitioners,

or more specific healthcare providers, such as neurologists, traumatologists, etc).

3. Paramedic describes the situation and symptoms through the phone; a remote (verbal) assessment is made, and the RHP makes recommendations on treatment and/or which healthcenter to transport the patient, so that they can receive the most relevant care.

The performance of this method is lacking in several ways. One of the most obvious ones is that the RHP is making an assessment based on the verbal descriptions of the paramedic, having no visual input from the situation. This leaves the RHP with only partial information which relies solely on the power of description of the paramedic. The obvious vulnerability in this system is one of the main reasons[2] Gjovik's Prehospital team has turned towards the development of more substantial tools such as PreViS.

2.2.1.2. Initiatives in Development.

The project owners for the PreVis system provided documentation of initiatives which preceded and are being developed parallel to it. These are based on the use of tablets embedded with proprietary live video-conference capability that is being used with their collaborating teams in Sweden. Much like the PreSSUB system, this program is also dedicated to stroke patients: The project is called "Video Support in the PreHospital Stroke Chain" (ViPHS). While extensive development and implementation has been managed by their teams, their documentation lacks evidence-based data gathering, thus it is very hard to understand if its performance is satisfactory in an Human Computer Interaction context.

In both previously mentioned projects, 'PreSSUB' and ViPHS, **there seems to be a lack of documentation regarding the observation and Human Factors / HCI considerations. It is concerning if said considerations are being monitored at all.**

As it has and will be seen throughout this report, this is a running theme; one which will surface once more when discussing the PreViS System.

[2] Another one of the main reasons why PreViS is relevant in Sweden and Norway has to do with the vast rural areas in which medical infrastructures are stretched too thin. This point will be described further in the following section.

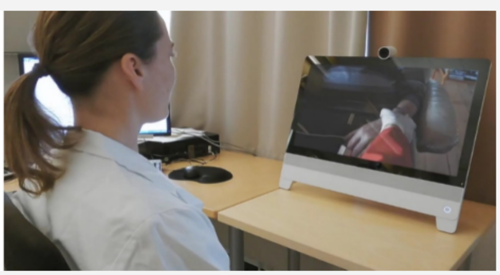
2.3. What is Previs?

The following section will serve as an overview of PreVis in the hopes of providing enough background to the reader so that the research decisions from the following chapters have been contextualized enough.

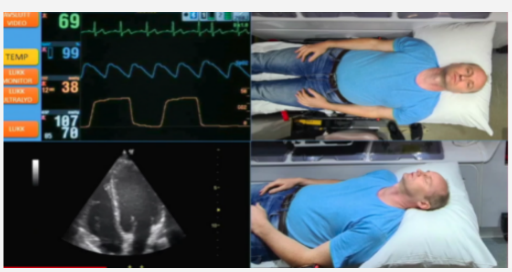
2.3.1. PreViS: In a Nutshell.

Previs is a pilot project which intends to integrate live video conferencing technology with the exercise of Paramedics. In a nutshell, Previs allows Hospital personnel to pre-visualize the state of a patient before they arrive at the hospital. The goal of this is to support the paramedic team to achieve early diagnosis and timely treatment of the patient. The following is a straight-forward example on how PreViS can be implemented in a potential real-life scenario.

Paramedics wear Headsets which can communicate live image and audio to a remote Health-Care Professional team in the hospital.



The hospital team are then able to support on rapid diagnosis and start treatment before the patient arrives to the Hospital.



The ambulances are equipped with cameras which livestream the status of the patient to the team that is in the hospital, this helps to monitor and treat the patient while the transportation is happening.

Fig. 2.02. 'Previs in a Nutshell'. [3]
An overview of the communication nodes of PreViS.
The photographs above belong to a simulated video created by the stakeholder to communicate their proof of concept. (Sykehuset Innlandet HF.2021)

[3] Complementing text and layouting is the authors' own, and was also featured in previous work. (Rivera, G., et al., 2021).

2.3.2. Description and Technical Details of Previs.

As it can be seen on the images above, PreViS is a system dependent on many interfaces, each one filled with its own set of complexities. For the purposes of this research, the scope covered is that which spans over the interactions of the Headset.

Fig. 2.03. Image from the Realware Headset which hosts the PreViS system. (RealWear, n.d.)

2.3.2.1. Headset Description.

2.3.2.1.1. Hardware.

The Headset hardware in question belongs to the American company RealWear. The device, called the HMT-1 is a wearable headband which integrates multiple functions destined to provide field workers with digital services and information. As it is mentioned on their website, the device was conceived to “instantly deliver information and IoT data to the frontline, including remote locations'' (RealWear, n.d.).

The headset holds various functions. The combination of an embedded video camera, video-call accessibility, internet access and a digital display, can provide multiple useful options for a worker practicing on unusual areas: *“[it can be used] for remote mentor video calling, document navigation, guided workflow, mobile forms and industrial IoT data visualization.”* (RealWear, n.d.). A large selling point of the configuration of the device is that it is conceived to be hands-free, having **all their points of interaction being voice-command based.**

For the purposes of this research it is relevant to describe some of the main components of the hardware. That being the mounted camera, and the digital display.

Fig. 2.04. Side view from the Realware Headset highlighting the position of the camera. (RealWear, n.d.)

2.3.2.1.1.1. Camera Hardware.

The device's camera is mounted on its side (their position can be reversed depending on user preference and/or hand or eye dominance).
The camera is not entirely fixed to the headset, being set on a pivoting Y axis which allows the user to adjust its angle.

Fig. 2.05. Posterior view from the Realware Headset highlighting the position of the microdisplay. (RealWear, n.d.)

2.3.2.1.1.2. MicroDisplay Specifications.

The Microdisplay is a 0,33 inch LCD screen which has an amplification crystal mounted on top. This surface magnifies the content of the small display, purportedly providing an image “like a 7" tablet”. On the microdisplay, the user can see the GUI of the system.
The display is the only digital visual point of interaction the user has with the device.

The display is mounted on a 'moving arm' or 'boom arm' which pivots on the Y axis and allows the user to adjust the microdisplay in front of their face.

2.3.2.1.1.3. The Visual Gulf.

The most common use-case for PreViS is to prompt the headset for a video call. In this scenario, if the user is interested in broadcasting their Point of View (POV), **they will receive a live broadcasting of what the camera is perceiving by looking at the microdisplay. In this way, they can see exactly what they are transmitting to the remote participant.** This highlights a particular experience of this headset, which is the gulf between the expectation of what the user thinks they are capturing with the camera, versus what the hardware of the camera is actually able to capture. To translate this into more quotidian terms, whenever a person sees a landscape and decides to take a picture of it with their mobile phone, there is always a difference between what the user sees with their eyes and what the camera is showing them on screen / able to pick up. In the case of the headset, this experience is replicated, albeit is further complicated by the fact that the very small display (0.84 cm diagonally) is always held limited inches away from their face, and some focusing and adjustment must be made by the user to understand this dynamic. This is further compounded by the fact that the main pivot of the camera is the users' neck and or head movements. From anecdotal first and second hand observation, it can be said that this can be a somewhat disorienting experience for users, at least on early attempts: users were often witnessed having to go through a period of adjustment — fidgeting, making trial and error adjustments, getting their head around the mental model of the interaction— when getting ready to use the device. This is a relevant issue to keep in mind, as it will be fairly prominent during this research.

Further, more granular technical specifications for the device can be found on the appendix.

2.3.2.1.2. Software.

The HMT-1 headwear can be used as a 'Plug and Play' device in the sense that developers can insert their own software on the system. This is the case of PreViS, where the software was recently developed by a Norwegian company named JodaPro.

2.3.2.1.2.1. What does the user see on the Microdisplay?

The following is a brief overview of the PreViS digital interfaces that are touched on a basic user journey. The images are pictures taken from the microdisplay itself. [4]

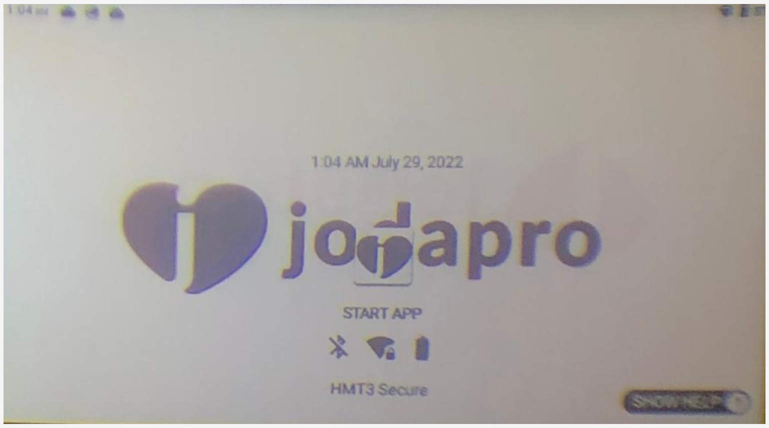


Fig. 2.06. Home Page UI
This is the first screen the user can interact with.

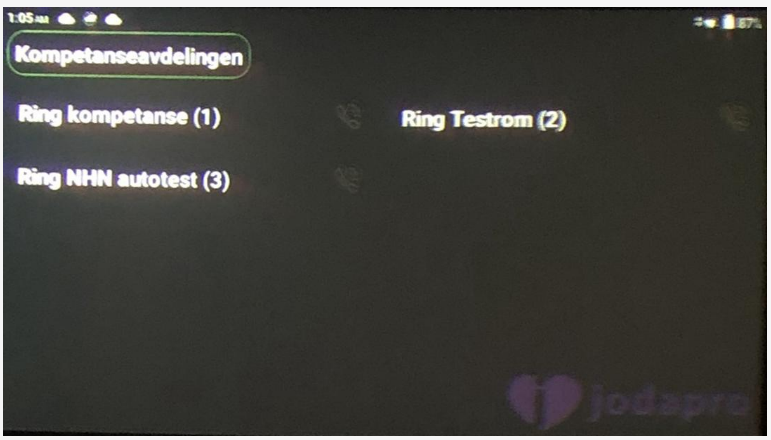


Fig. 2.07. 'Kompetanseavdelingen' UI.

This takes the user to the Kompetanseavdelingen section of the system. This opens the call log screen where users can select the number they want to video-call to. Voice Command: Start App

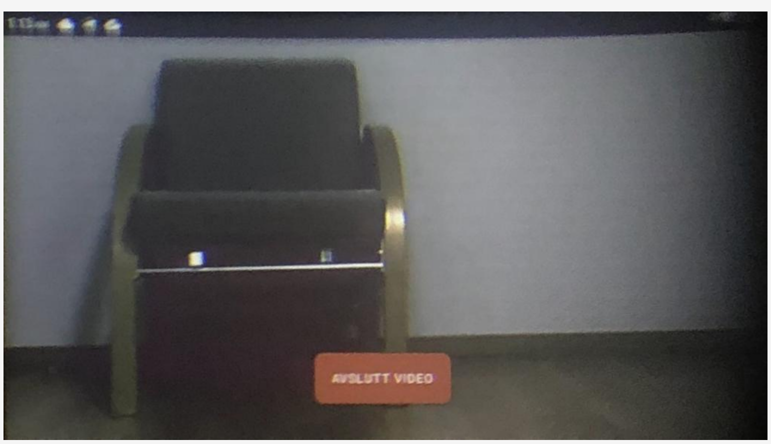


Fig. 2.08' Call Space' UI.

Users arrive at the 'Call Space' where video calls happen (Voice Command: Ring [Selection). In this section the microdisplay shows what the camera is broadcasting, and a button which says 'Avslutt Video', which allows the user to terminate the call. In this space users can use the commands: Zoom, and Light to trigger enhancement on the camera, and the ignition of the integrated LED flashlight, respectively.

---

[4] The low quality of the images is due to the difficulty of capturing the small 0,33 inch display on camera. Because of the configuration of the same (having a mounted magnifying glass on top), the edges of the UI appear blurry unless the user tilts their view to focus on it.

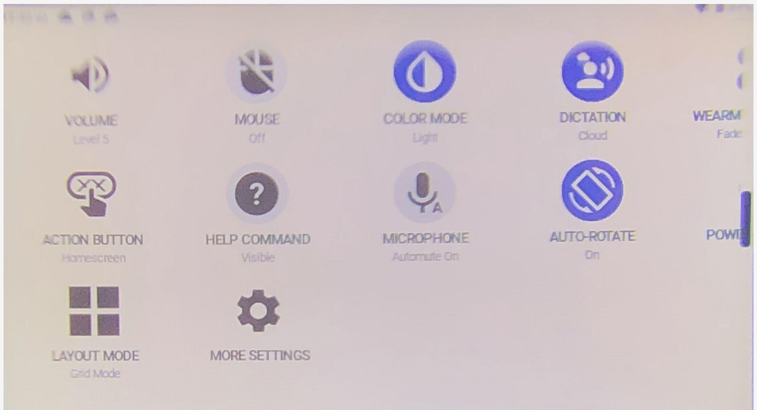


Fig. 2.09. 'Configuration' UI.

Section with configuration tools (Voice Command: My Controls).

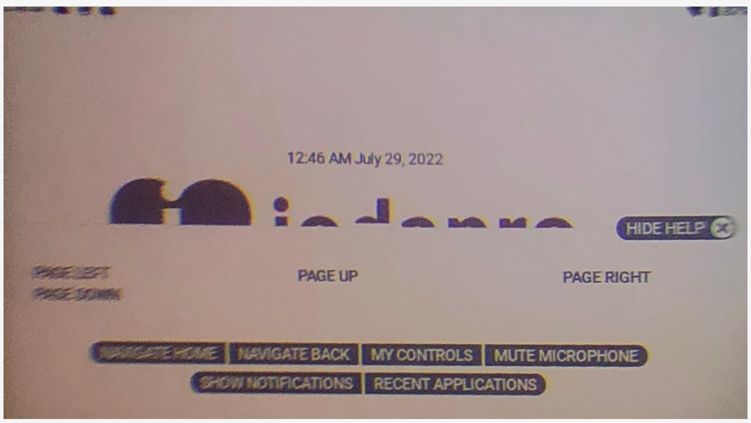


Fig. 2.10. 'Show Help' UI

This shows a menu with different options (Voice Command: Show Help).

2.3.3. The PM-RHP Interaction:

The current way PreviS is being used is as follows:

1. The Paramedic (PM) calls the Remote Healthcare Professional (RHP) from a mobile phone in the same way they would outside of the PreViS implementation.
2. The PM will explain the basics of the emergency and request the RHP if they wish to see the patient using PreViS in a specific location (for example, the screens on the trauma room).
3. After agreeing on these details, the PM can call the specified location with PreViS, and the RHP will receive the PM's video feed on the monitors.
4. The RHP can see the patient and initiate the medical assessment.

2.4. Previous Research: General Context of Previs.

The following section discusses knowledge acquired which preceded the current research. These observations were the early inspiration / justification for this thesis.

2.4.1. Is the potential of PreViS valuable for Pre-hospital practices?

It was important for this research process to understand if the implementation of a system like this would have realistic potential to contribute positively towards the practice of prehospital services. In other words: **Would this system truly respond to relevant user needs?**

2.4.1.1. Previous Research: Feedback from Pre-Hospital Professionals (PHPs)

During the autumn of 2021, as part of a separate research, the author conducted several first-hand interviews with different members of the prehospital chain (paramedics, a trauma-coordinator and a veteran paramedic trainer) both inside and outside of the PreViS project management teams. In this process it was discovered that PreViS garnered overwhelmingly good opinions on behalf of Pre-Hospital Professionals, whether they had used it or not (Rivera, G. et al., 2021). There was a perception that PreVis could provide a very welcomed sigh of relief to some of the most acute pressure points and stressors in their practice.

The areas of 'relief' that were most highlighted by the interviewees were:

- PreViS can provide decision support for paramedics attending to calls of high complexity, allowing for life-saving treatment to be administered at the most needed time. This is especially relevant in cases such as stroke, complex trauma, pregnancy, et al.

- By providing a live video of incoming emergency cases, hospital personnel can pre-visualize the condition of the patient, and thus make logistical preparations in terms of manpower, team coordination, equipment preparation, etc.

- The Rural Context: This project is being championed by regions of Norway and Sweden which are very sparsely populated and set among large geographical distances. These conditions make it very hard to deliver equal quality of healthcare to all inhabitants.

According to the personnel that were interviewed from Gjøvik Sykehuset and Lillehamer Sykehuset, for some emergency calls that happen in rural environments in Norway, it can take up to four hours to arrive at a health center. In these scenarios PreVis can be a literal lifesaver. By having live streaming video and audio, a Healthcare Professional in the hospital can provide specific decision support to the paramedic in the ambulance throughout an extended period of transportation.

2.4.1.2 Feedback from Scientific Literature

While the implementation of systems similar to PreViS are quite new, and thus not a lot of scientific research has been published on the matter, tangential research regarding prehospital factors could possibly indicate that a tool which provides the PHP with timely decision support could possibly prevent adverse consequences for the patient.

Literature such as the *Emergency Medicine Journal* publication 'What causes adverse events in prehospital care? A human-factors approach' indicates that disconcerting patient factors may lead to uncertainty from the PHP. This consequently could lead to procedural omission in treatment, and thus patient harm events may occur (Price et al., 2012). It is possible to theorize that by adding live decision support from a remote Healthcare professional, PHP uncertainty would be mitigated, and thus possible patient errors could be avoided .

Disconcerting patient factors → paramedic uncertainty → clinical omissions → serious patient harm events.

Fig. 2.11. Chain of factors that lead to serious patient harm events according to Price et. al (2012).

2.4.2. The State of PreViS' Product Development.

The Project is currently doing a fashion of soft launching. As proof of concept, there is currently one ambulance suited with the equipment. However, at the date of this writing, there are dozens of Headsets being distributed around the paramedic institutions of Innlandett, and the product owners are using **medical simulations exercises** or **roleplay** in order to train personnel on how to functionally use the system.

2.4.2.2. Pseudo-User Testing: The Role of Roleplay on PreViS' R&D.

The current user testing done for this project is based primarily on feedback obtained during the aforementioned sessions of roleplay. The following is an explanation on the status and tools used for user testing the current iteration of PreViS. **This discussion allows understanding in the ways the current methodology is failing, laying the foundation for the methodological design of the current research.**

During the autumn of 2021, the researcher interviewed experienced paramedic trainers in the methods that are being utilized to train paramedics on the use of PreViS. The most implemented teaching method was Roleplay: This is a traditional method in the healthcare environment used to simulate real-life scenarios professionals might encounter in their work. This allows them to learn and practice strategies and methods within a safe space (Nestel, D., 2007) .

The methodology of the implemented simulations were described during interviews as follows:

1. The trainer gathers the group of paramedics, and distributes the roles (i.e. someone plays the patient, the acting paramedic team, and bystanders at the scene).

2. The trainer decides the scenario that will play out (i.e what type of emergency are they simulating[5]).

3. They 'play-out' the scene.

4. After the scene ends, the team gathers and verbally provides feedback to each other on what happened during the exercise. This stage is traditionally known as 'debriefing'. This type of debriefing can last five to ten minutes, and then another roleplay would be acted out, often with participants switching roles.

[5] The person playing the patient is often given a brief in private as to the pathology they should simulate, in order for the other paramedics to react in real time to the displayed symptoms.

As referenced before, **these roleplays and the feedback from the debriefing sessions are the closest to comprehensive user-testing as the project has implemented so far**.
This method of user testing is quite deficient for multiple reasons:

- The improvised nature of the roleplays that were observed lack structure and the intentionality of a true user test. There is no consistency from one roleplay to another: variables (scenarios, conditions and types of users) can differ widely from one roleplay to another. This can make the data acquired during these ‘user tests’ inconsistent and thus hard to analyze.

- The debriefing from the ‘users’ is verbal, public and quick: Debriefings happen briefly, right after roleplays, by speaking in front of everyone in the team. This type of debriefing is problematic in the context of user-testing for multiple reasons, some of which are summarized below:

    i) Public speaking is not the optimal way to get honest feedback from a user. There is little time for thought, pressure to speak ‘on the fly’. And ‘loudest voices’ may derail the conversation from important feedback to be expressed (Stickdorn, M. 2018).

    ii) There might be invisible power dynamics at play which might prevent honest feedback from users, For example, one of the main trainers from the PreVis process is an experienced member of the local paramedic unit, who is also one of the Project Owners. It is reasonable to believe that some users may hesitate to point out important flaws in the PreViS system due to conscious or unconscious biases or social power dynamics.

    iii) The data from the ‘user-tests’ is not being documented. There are no written records on usability related-findings, thus analysis can’t be performed in an official capacity.

- This amalgamation of training and user-testing might muddy the waters in the way of obtaining critical feedback. As an outsider observer, the researcher found it hard to interpret whether participants know that they are allowed to criticize the system, or whether their role is to *train* for it (e.g. adjust to it and '*make it work*').

- The Paramedic Silo: During the previous phase of this research, which was conducted through a service design project, the researchers found that externals to the paramedic circle are usually not included in the roleplay (Rivera, G., et.al., 2021). This means that the role of the remote healthcare professional, i.e, the person on the other side of PreViS, providing the decision support in the hospital (usually a General Practitioner doctor or a specialist) is currently exclusively being 'played' by a paramedic. In this way, the remote healthcare professional's roles, wants and needs are being interpreted by a paramedic who may or may not be aware of them[6]. By excluding this user base from directly participating in these simulations/pseudo-user-testing, the process probably is missing important insight and possible unidentified relevant pain points that might hinder the overall performance and impact of the PreViS system.

As previously expressed, the current state of launching is such that training is the priority, and user-test is a byproduct. This endangers the process of user-testing and understanding how usable the system really is.

2.5 Lack of Usability Studies.

As described in the previous section, during several interviews conducted with the project managers, and during extended observation, it was revealed that despite the tremendous technical development that has been conducted by their team, **there is no HCI or Usability evaluation process to understand how human factors may influence the use of the system.**

From an interaction design perspective, this was quite an alarming find, as a system that functions in such a complex environment should be thoroughly user tested. Unlike a casual app, this system will function in a context where there is very little room for error, as malfunctions can affect the cognitive load and stress levels for Pre-Hospital Professionals (PHPs) in critical situations. As previously mentioned, studies about PHPs have determined that stress can have a significant and quantifiable effect on their cognitive abilities (Summers, A. 2010). Because of

[6] On all first-hand observations this role was always played by a senior Paramedic, and/or Project Manager of PreVis.

this, the introduction of PreVis on their every-day workings must provide ease of use and guarantee little to no adverse factors on its users, as these could possibly result in affected performance and possible adverse events on patients.

In order to identify any possible factors that might impact the PHP, extensive user testing in realistic environments should be carried out. Due to circumstances related to project development, and scheduling constraints related to the launching, this user testing[7] would have to be done with the participation of real PHPs, adhering to realistic circumstances, and with an emphasis on observing usability and HCI. This motivated an elaborate methodology, experiment and evaluation design which could satisfy these variables. The steps taken and variables observed to satisfy these requirements will be described in the following chapter.

[7] Previous works the author has conducted as part of a team dealt with reshaping the roleplay as a more efficient way of user-test (Rivera, G., Sæbø, S., et al., 2021). However, due to time restrictions from the project, this route may not be viable.

3. Methodology.

3.1. Research Design.

In general terms, the purpose of this research was to capture how the PreViS system helps actors interact with each other in Prehospital emergency scenarios. These are deeply complex environments with multiple unique challenges, in which adding this new variable may alter the interaction in unexpected, subtle ways, ways that may not be able to be captured through interviews or other more traditional mediums.

An ideal approach for this case would be to implement field studies and observations. However, in the context of paramedic emergencies dealing with real patients, research exercises can bring with themselves a set of critical challenges (patient privacy rights for one), which may make the implementation of the research difficult to coordinate (Söderholm et al., 2007), impractical, or even unattainable in the timeframe of this research.

Because of this, the research design of this project turned to a proxy of natural observation, which is the use of controlled simulations. The use of simulations, also known as roleplay, is already inherent in paramedic practice, as it is one of the main methods for training, thus participants would be familiar with it and not intimidated by it.

3.1.1. What is Roleplay?

In a general context, roleplay can be defined as: *"(...) an instructional method where learners take on the responsibility of representing different character roles, within predefined, often realistic scenarios"* (Bawa, 2020). Van Ments describes it as follows:

The idea of role-play, in its simplest form, is that of asking someone to imagine that they are either themselves or another person in a particular situation. They are then asked to behave exactly as they feel that person would. As a result of doing this they, or the rest of the class, or both, will learn something about the person and/or situation. In essence, each player acts as part of the social environment of the others and provides a framework in which they can test out their repertoire of behaviors or study the interacting behavior of the group (Van Ments M. 1989) (Nestel, D., Tierney, T. 2007).

In the medical environment, simulations are widely used as a training method where learners can practice techniques and procedures in a safe space. An interesting aspect of this method is that, as mentioned in the quote above, simulations allow for the observation of the interacting behavior between the involved parties. In the prehospital environment, for example, these types of simulations often include one member of the paramedics playing the part of the patient, another playing the part of the caretaker, dispatcher, etc. After the simulations are done there is often a debrief with each party describing what they observed during the interaction. At the end of each ‘play’, there is a vast cache of information which can be gathered, either from capturing and chronicling of the events that occurred —performance, interactions— (external observations), and the feedback from the participants (internal observations).

There are multiple ways to implement a simulation: They can be fully scripted, where players follow pre-written interactions exclusively, or partially scripted, where players only follow a limited list of prompts (Nestel, D., Tierney, T. 2007). Other variations include a structure where some of the players are following a script, while other players are simply reacting to it: for example, a scripted player portraying the patient, whose task is to depict a specific symptomatology for the unscripted players to figure out or react to. Henceforth this mode will be referred to as semi-blind simulation.

Medical simulations such as the ones described above have been observed to be used in Norway in the prehospital environment.

3.1.2. Related Research.

Simulations can also be used for studying the functionality of medical devices and telemedicine. One example of this proved to be very relevant to this research. Such is the case of the research paper “*The potential impact of 3d telepresence technology on task performance in emergency trauma care*” by Hanna M. Söderholm, et.al (2007), where their goal was to explore the potential of 3D telepresence in the case of pre-hospital emergencies. They used a semi-blind simulation setting where one control and one intervention group had to use different mediums (2D audiovisual transmission and a 3D telepresence proxy, respectively), to conduct a diagnosis and procedure with the support of a remote healthcare professional.

“(...) We conducted an experimental evaluation that compared collaboration between paramedics and remote physicians under two conditions. One condition utilized high quality, state-of-the-art 2D videoconferencing to support collaboration during a simulated emergency medical situation. The other condition utilized a 3D telepresence proxy because [...], the 3D telepresence technology does not yet exist. In each condition the same emergency medical situation was simulated. A professional paramedic was asked to diagnose and treat a trauma victim, and a physician was available to assist the paramedic. In the 2D condition, the paramedic and physician interacted via state-of-the-art bidirectional and remote-controlled 2D video and audio. In the 3D proxy condition, the paramedic and physician were collocated but the physician could not touch anyone or anything in the room. Performance and posttest data were collected and analyzed to compare the impact of technology in each condition.” (Söderholm et al., 2007 p.81).

Amongst the data captured by the researchers in the same paper were ‘**harmful interventions**’. these were defined as “*treatments or procedures that are potentially harmful for the patient by increasing risks for complications or unwanted side-effects and by delaying the time it takes to the [patient’s to reach a stable condition]”*[8] (Söderholm et al., 2007 p.83).

As it will be seen in the following sections, this experiment served as the main reference for the creation of the main experiment in this research.

3.2. Designing the Simulation.

In general terms, what the designed simulation expected to achieve was to **observe how paramedics interact with the PreViS system in order to get remote guidance and decision support throughout an unfamiliar patient case.** The following section will describe the pathway to the crafting of the experiment.

[8] This is a term and definition which will be adapted into the evaluation framework of this research.

3.2.1. Selecting the Emergency.

Following Söderholm's experiment framework, it became important to select an appropriate medical emergency to simulate. The choice would have to possess some relevant criteria for the purposes of this research. The emergency selected would have to:

1. Require the paramedics to perform a structured sequence of examinations, in this particular case, one which they would be unfamiliar with, and would need to depend on remotely given instructions. This fulfilled two purposes:

    a. Allow the observation of how an RHP can remotely guide a paramedic through a complex examination in both control and intervention groups.
    b. A structured sequence of examinations allows easy replication and consequent comparison between simulations and participant groups. It also makes it easy to verify if the examination has been fulfilled correctly and to completion.

2. Have measurable outcomes.
3. Have relevance in the context of what the local Norwegian paramedics are experiencing in their exercise.

Söderholms et.al. (2007) experiment chose as an emergency a symptomatology which would require the implementation of a cricothyrotomy and the management of a difficult airway. When cross checking the design of the experiment with local paramedics and experts, the feedback was that such scenario is not common nor expected from the prehospital practice in Norway[9]. Because of the above, the selected scenario to simulate became a patient presenting symptomatology consistent with a posterior stroke. This selection permitted the following:

- Allowed the paramedic to confront a patient with a confounding symptomatology, i.e. a scenario where they would typically need decision support.
- A posterior stroke is a condition which can be diagnosed through a sequence of examinations called NIHSS.

[9] The reference experiment took place in the United States. The expanse of what paramedics are legally allowed or required to do in emergency cases varies widely to what is practiced in Norway.

Definition of NIHSS

NIHSS stands for National Institute of Health Stroke Scale. It is an assessment tool created to evaluate patients who may be experiencing acute stroke.

“The NIHSS is a 15-item neurologic examination stroke scale used to evaluate the effect of acute cerebral infarction on the levels of consciousness, language, neglect, visual-field loss, extraocular movement, motor strength, ataxia, dysarthria, and sensory loss.

A trained observer rates the patient’s ability to answer questions and perform activities. Ratings for each item are scored with 3 to 5 grades with 0 as normal, and there is an allowance for untestable items. The single patient assessment requires less than 10 minutes to complete.

NIH stroke scale

Admission date:

Item	Score	Description	Side	Before	2 h	24 h	7D/ Disch
			Time:				
1a.Level of consciousness	0	Alert					
	1	Not alert, but arousable with minimal stimulation					
	2	Not alert, requires repeated stimulation to attend					
	3	Coma					
1b. LOC questions *Ask patient the month and their age*	0	Answers both correctly					
	1	Answers one correctly					
	2	Both incorrect					
1c. LOC commands *Ask patient to open/close eyes and form/release fist*	0	Obeys both correctly					
	1	Obeys one correctly					
	2	Both incorrect					
2.Best gaze *Only horisontal eye movement*	0	Normalt					
	1	Partial gaze palsy					
	2	Forced gaze palsy					
3. Visual field testing	0	No visual field loss					
	1	Partial hemianopia					
	2	Complete hemianopia					
	3	Bilateral hemianopia (blind, incl. Cortical blindness)					
4. Facial palsy *Ask patient to show teeth or raise eyebrows and close eyes tightly*	0	Normal symmetrical movemeent					
	1	Minor paralysis (flattened nasolabial fold, asymmetry on smiling)					
	2	Partial paralysis (total or near total paralysis of lower face)					
	3	Complete paralysis of one or both sides (abscence of facial movement in the upper and lower face)					
5. Motor function arm	0	Normal (extends arm 90° or 45° for 10 sec without drift)	**Right**				
	1	Drift					
	2	Some effort against gravity					
	3	No effort against gravity	**Left**				
	4	No movement					
	9	Untestable (joint fused/limb amputated) *(do not add score)*					
6. Motor function leg	0	Normal (holds leg in 30° position for 5 sec without drift)	**Right**				
	1	Drift					
	2	some effort against gravity					
	3	No effort against gravity	**Left**				
	4	No movement					
	9	Untestable (joint fused/limb amputated) *(do not add score)*					
7. Limb ataxia	0	No ataxia					
	1	Present in one limb					
	2	Present in two limbs					
8. Sensory *Use pinprick to test arms, legs, trunk and face, compare side to side*	0	Normal					
	1	Mild to moderate decrease in sensation					
	2	Severe to total sensory loss					
9. Best language *Ask patient to describe picture, name items*	0	No aphasia					
	1	Mild to moderate aphasia					
	2	Severe aphasia					
	3	Mute					
10. Dysarthria *Ask patient to read several words*	0	Normal articulation					
	1	Mild to moderate slurring of words					
	2	Near unintelligible or unable to speak					
	9	Intubated or other physical barrier *(do not add score)*					
11. Extinction and inattention *Use visual double stimulation or sensory double stimulation*	0	Normal					
	1	Inattention or extinction to bilateral simultaneous stimulation in one of the sensory modalities					
	2	Hemi-inattention, severe or to more than one modality					
12. Distal motor function *Ask patient to extend his/her fingers as much as possible*	0	Normal	**Right**				
	1	At least some extension after 5 sec but not fully extended					
	2	No voluntary extention after 5 sec	**Left**				
		Total score:					

Fig. 3.01. NIHSS Scale form which details the steps of the examination. For ease of reading, a larger version of this image can be found on the appendix. From www.nihstrokescale.org, (n.d.)

The evaluation of stroke severity depends upon the ability of the observer to accurately and consistently assess the patient*”* *(www.nihstrokescale.org, n.d.).*

Physicians will go through the 15 item list through examination and observation to see how the patient responds to their specific stimuli. For example, to test NIHSS item #4: Facial Palsy, the patient is asked to show their teeth, raise their eyebrows, frown, and slightly close their eyes. A list of the prompts to the patients and expected responses can be seen in Fig. 3.01.
The inclusion of NIHSS in the simulation fulfilled the requirement for having a predetermined sequence of specific examinations paramedics would have very little to no familiarity with, which can be easily comparable and evaluated for success or failure, and which is relevant to the norwegian context.

3.3. Setup of the Simulation.

The setup of the scenario, symptomatology and script was developed in collaboration with prehospital professionals who provided their experience and feedback on the matter.

3.3.1. Simulated Task.

As previously mentioned, in a simulated scenario the participating paramedics were asked to conduct a relatively complex examination process they had not practiced before (NIHSS examinations). The execution of such was assisted by the guidance of a Remote Healthcare Professional (RHP). In the control group they conducted this process over a mobile phone call—representing the status quo of the praxis in Norway— while in the intervention group they used the PreVis system instead.

3.4. Simulation Details.

The objective for the participants was to conduct the NIHSS examination correctly and to completion with the guidance of the RHP. Through this examination the paramedic (henceforth denoted as PM) had to report and/or capture the symptoms of the person playing the patient. These were as follows:

- Left eye palsy (discoverable on NIHSS step #2).
- Compromised Visual Field (discoverable on NIHSS step #3).

All other stages of the NIHSS examination should have been 'cleared' without any other symptomatology discovered.

These symptoms paint the picture of a patient experiencing a posterior stroke. The correct course of action with a patient of this symptomatology would be to be transported as soon as possible to a health center that is equipped with an MRI machine. Failing to diagnose the posterior stroke in this scenario may result in the transportation of the patient to a different health center where they may experience further delays on getting time-sensitive possibly life-saving treatment.

3.4.1. Setting the Scene.

In the following section the role of each of the participants will be described. Following that, a sequence of the designed path of the simulation will be explained.

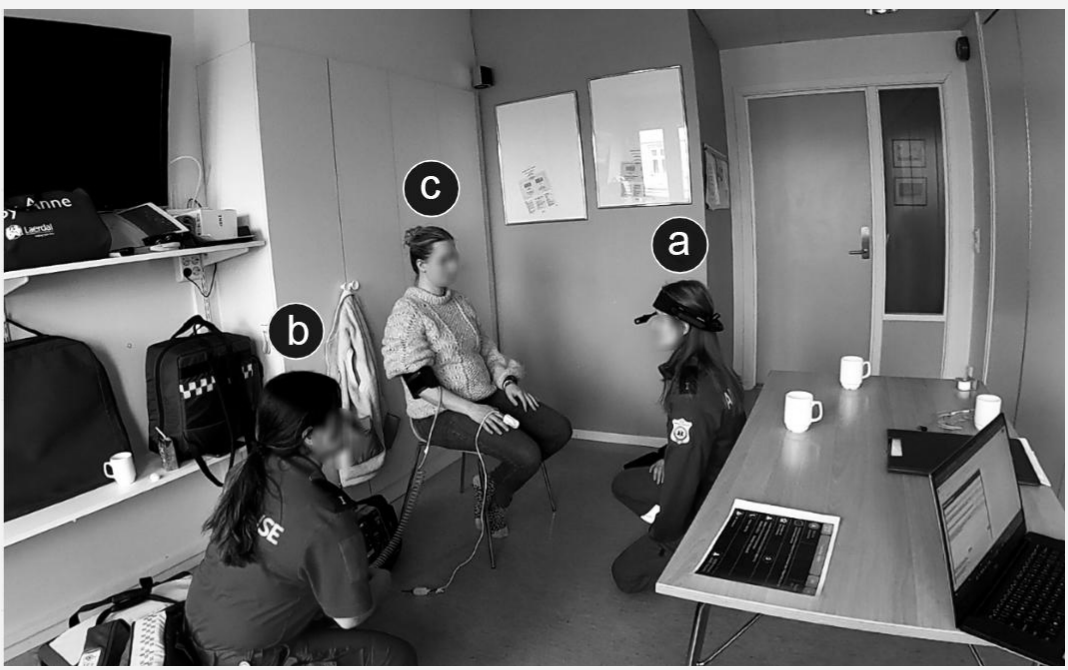


Fig. 3.02. Configuration of the simulation. Pictured: a) Primary PM; b) Partner PM; c) Patient.
Not pictured, RHP communicating remotely on the device; observers, standing behind the camera.

3.4.2. The Participants.

a) **Primary PM:** This role is played by the participants of the experiment. They are the focus point of the observations. This is a non-scripted part: They are the more 'reactive' player, as neither the topic nor expected sequence of the simulation has been disclosed to them. It is their role to play the scenario as it unfolds, and react in real time to the events.

b) **Partner PM:** A semi-scripted part: They were briefed on the expected sequence of the roleplay and asked to only support the Primary PM in passive ways: Being a support in discussions but handing the first prerogative to the Primary PM. They were not briefed on what would be the patient's symptomatology. They were played by a paramedic of medium experience.

c) **The Patient:** A scripted part: They were carefully briefed to display a particular symptomatology which would signal the posterior stroke diagnosis. They were played by a healthcare professional (nurse), with multiple years of experience in first-hand treatment of patients.

d) **The RHP**: A semi-scripted part: An off-site healthcare professional (paramedic), was standing by to receive the phone call from the PM. They were familiar with the concept of PreViS and had tested the device on a previous occasion. They were not briefed on what the patient's symptomatology would be. However, they were clued into the fact that they would need to perform NIHSS examinations, and were given the necessary documentation about it. They were asked to avoid large deviations from the interaction from one participant to another.

e) **The Observers:** Non-active parts: Two observers were in the room. One was the project owner (A paramedic trainer with decades of experience). The other was the researcher. The latter was there for first hand observations, as well as to monitor the cameras that were recording the simulation. Neither observer interacted with the paramedic during the simulation. This was only breached to signal the end of the same.

3.4.3. Simulation Sequence.

1. **Preparation:** The participants were allowed 20 minutes to practice the basic interactions with the headsets by themselves. They were introduced to the PM partner, and waited for the beginning of the simulation.

2. **Beginning of the Simulation:** Set in a closed room, the Primary and Partner PM were given a call through radio simulating a true call from the Emergency Dispatch. After this, they were provided a piece of paper which replicated basic emergency-related information they typically receive while driving on the ambulance. They were then given 5 minutes of ‘preparation time’, as a way to replicate a version of the interaction paramedics perform with each other while on-route to the patient’s location. In this space, paramedics usually discuss the initial strategy they will enact when reaching their destination.

3. **Arrival at the Scene:** Here, both paramedics entered the room where the emergency was being simulated. Then, the Primary PM conducted the initial parts of the examination (ABC routine). Following this, the Primary PM briefly discussed the course of action with their partner. In this case, the patient was simulating signs which are either confounding to the PM or signaling the possibility of stroke. In either case, the typical course of action was to call the RHP and get decision support remotely. **For the control group this involved calling the RHP on a mobile phone and making the examination over the call. For the intervention group this meant calling the RHP with a mobile phone, agreeing to connect through PreVis, then connecting, and conducting the examination through the headset.**

Fig. 3.03. Configuration of the simulation.

4. **Examination with RHP:** In this phase, the Primary PM debriefed the RHP about their current findings. Following that, the RHP instructed the PM how to conduct NIHSS on the patient. When completing the examination, the RHP gave their recommendation on where to transport the patient.

5. **End of Simulation:** The Observer (paramedic trainer) announced the end of the simulation once the call with the RHP was over.

3.4.4. Post-Simulation Data Gathering.

Participants (PMs and RHP) were given post-test questionnaires. The PMs were given a paper-based System Usability Scale (SUS) questionnaire which was filled in private and anonymously, in pursuit of honest feedback (this will be addressed with more detail in subsequent sections).

The RHP was given a digital bespoke set of questions, relevant to their experience (see Feedback from RHP section). After each round of simulation the researcher interviewed both the trainer paramedic, and the healthcare professional who played the patient in order to capture their observations on what happened in the role-play.

3.5. Simulation Evaluation Structure.

The purpose of the simulation was to capture a cache of information about the interaction of the parties with the system. After this was achieved, the data was processed through an evaluation framework and a usability inspection.

3.5.1 Usability Inspection: Performance Evaluation Frameworks.

According to Jakob Nielsen, an Usability Inspection is defined as a blanket term for the exercise of evaluators implementing a set of methods in pursuit to evaluate the usability of an user interface: (...) *normally, usability inspection is intended as a way of evaluating user interface designs to find usability problems. In usability inspection, the evaluation of the user interface is based on the considered judgment of the inspector(s). The individual inspection methods vary as to how this judgment is derived and on what evaluative criteria inspectors are expected to base their judgments".* (Nielsen, J. 1995).

In this case, the usability inspection of PreViS will take the findings from the simulation observations, sift them through a framework of inquiry, and then evaluate them through a lens of a heuristic evaluation. This structure of evaluation will be explained in the following pages.

3.5.1.1. Observation Guide and Objectives of the Simulation.

The first step of this process was to clarify what were the main objectives and lines of inquiry of the simulation.

The following is a list of the observation targets which will be pursued during the simulation. These derive directly from Q1 and Q2.

Q1: Can the system help make examinations more accurate in comparison to the control group?

Research Question	Performance measure	Performance measuring tools
Q1.1. Can complex examinations be performed more accurately in comparison to the control group?	Was the accuracy of the examination higher when done through PreVis?	● Performance Evaluation Framework (see next section). ● Video footage inspected by the researcher and an experienced paramedic trainer.
Q1.2. Can the system help avoid patient harm?	Are there fewer harmful interventions with PreVis, compared to the control group ?	● Comparison of harmful interventions between both groups. ● Performance Evaluation: Framework. ● Video footage evaluated by the researcher and an experienced paramedic trainer.
Q1.3. Can the symptomatology of the patient be effectively transmitted to the RHP?	Can all signs and symptoms that the patient is presenting be captured on video? Is the audio and video quality of the device high enough for the RHP to make an educated decision?	● Inspection of footage and identification of signs captured and/or lack thereof. ● Video footage evaluated by the researcher and an experienced paramedic trainer. ● Posttest questionnaire from RHP participant on the simulations.

Table. 3.01. Table showing how the research questions of Q1 will be evaluated and analyzed.

Q2: Is the Usability of the system acceptable for its use on PreHospital Emergencies?

Matter of Observation	Performance measure	Performance measuring tools
Q2.1. Do users perceive the system to be easy to use?	Feedback from users.	• Posttest Questionnaires System Usability Scale (SUS) Survey. • Performance Evaluation Framework. • Review of footage observing behavior: user errors, system failures. • Interaction/Navigation analysis (Instone Test). • Performance Evaluation: Framework (see next section). • Usability Inspection: Device/system-wide Heuristic evaluation.
Q2.2. Is the use of the device causing an extra load (cognitive or otherwise) on the user.	Does the setup of the device respond to the user's mental model? Are there consistent user errors?	
Q2.3. Is the HCI intuitive? (Voice-command interaction, navigation etc).	Are the touchpoints of the interaction intuitive? Does the language used in the system (Voice commands, sections labeling), match real world conventions. Is the system's navigation clear and accessible to the average user.	
Q2.4. Is the interface prone to user errors?	Does the interface violate usability heuristics?	
Q2.5. Is the system prone to technical errors?	Are there consistent malfunctions in the system?	

Table. 3.02. Table showing how the research questions of Q2 will be evaluated and analyzed.

3.1.5.2. Performance Evaluation Frameworks.

Due to the live and somewhat fleeting nature of roleplays, observation goals and clear evaluation frameworks must be grounded so that objective, comparable information can be distilled from the exercises. With this in mind, **in order to answer Q1 and its sub questions in an objective way, a Performance Evaluation Framework was created**. The creation of this artifact **allowed a structured and comparable observation process between simulations, which would give enough data to objectively answer the research questions.**

As a reminder, the aforementioned inquiries of Q1 are as follows:

Q1: Can the system help make examinations better in comparison to the control group?
● Was the accuracy of the examination higher when done through PreVis? ● Are there fewer harmful interventions with PreVis, compared to the control group? ● Can all signs and symptoms that the patient is presenting be captured on video? ● Is the audio and video quality of the device high enough for the RHP to make an educated decision?

After each simulation, the events were chronicled in an overview, and then interrogated by the questions of the framework. Said questions aimed to give enough background to answer Q1 and Q2 more thoroughly.

Some of the inquiries interrogated were the following:

- Was the examination efficient?[10]
- Could NIHSS be practiced correctly?
- Were there harmful interventions?[11] If so, how harmful were they?

---

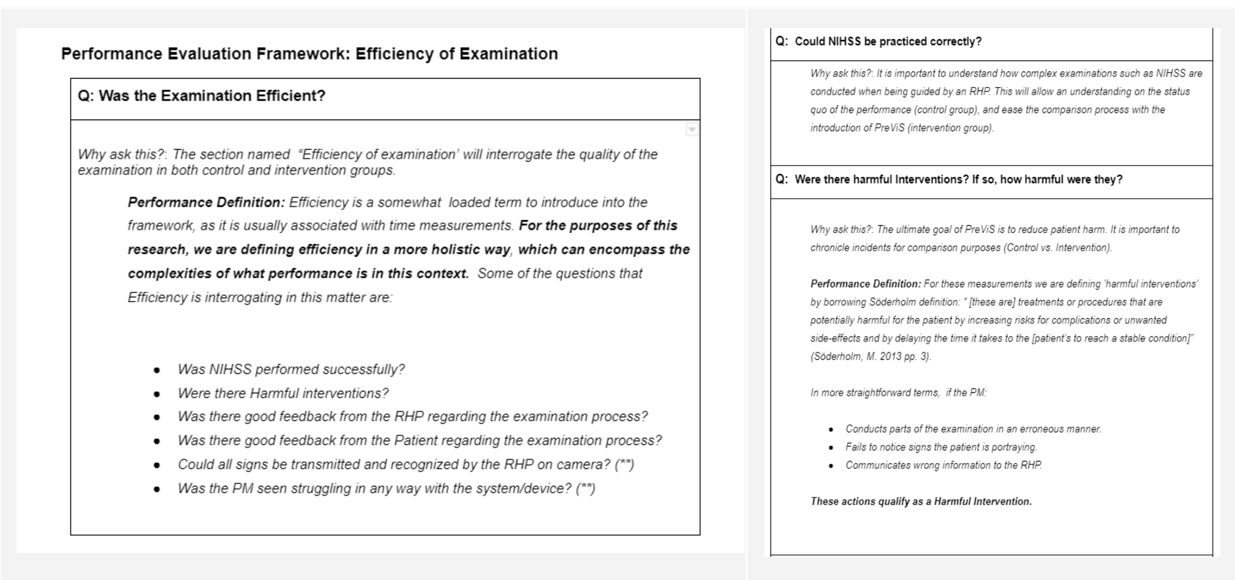
[10] The metrics on what is defined as efficiency in this context is detailed in the appendix.
[11] Harmful Interventions is a term 'borrowed' from Söderholm's experiment. For a complete definition of the term and metrics, see the appendix.

- Were there any instances of user error?

A performance evaluation framework was also created to capture the Feedback from the RHP. Some of the questions asked of the RHP were as follows:

- How easy was it to conduct the examination through the phone/headset?
- Was there ever a drop in quality of video or audio that got in the way of examinations?
- Was the image of the patient in focus enough so that you could see what you needed from the patient?
- How confident are you that you got all the information you needed to get a diagnosis?
- How confident are you that the examination from the paramedic was done properly?

A detailed, question by question breakdown of the frameworks, accompanied by the rationale behind every inquiry, is included in the Appendix.



Performance Evaluation Framework: Efficiency of Examination

Q: Was the Examination Efficient?

Why ask this?: The section named "Efficiency of examination' will interrogate the quality of the examination in both control and intervention groups.

Performance Definition: *Efficiency is a somewhat loaded term to introduce into the framework, as it is usually associated with time measurements.* ***For the purposes of this research, we are defining efficiency in a more holistic way****, which can encompass the* ***complexities of what performance is in this context.*** *Some of the questions that Efficiency is interrogating in this matter are:*

- *Was NIHSS performed successfully?*
- *Were there Harmful interventions?*
- *Was there good feedback from the RHP regarding the examination process?*
- *Was there good feedback from the Patient regarding the examination process?*
- *Could all signs be transmitted and recognized by the RHP on camera? (**)*
- *Was the PM seen struggling in any way with the system/device? (**)*

Q: Could NIHSS be practiced correctly?

Why ask this?: It is important to understand how complex examinations such as NIHSS are conducted when being guided by an RHP. This will allow an understanding on the status quo of the performance (control group), and ease the comparison process with the introduction of PreViS (intervention group).

Q: Were there harmful Interventions? If so, how harmful were they?

Why ask this?: The ultimate goal of PreViS is to reduce patient harm. It is important to chronicle incidents for comparison purposes (Control vs. Intervention).

Performance Definition: *For these measurements we are defining 'harmful interventions' by borrowing Söderholm definition: " [these are] treatments or procedures that are potentially harmful for the patient by increasing risks for complications or unwanted side-effects and by delaying the time it takes to the [patient's to reach a stable condition]" (Söderholm, M. 2013 pp. 3).*

In more straightforward terms, if the PM:

- *Conducts parts of the examination in an erroneous manner.*
- *Fails to notice signs the patient is portraying.*
- *Communicates wrong information to the RHP.*

These actions qualify as a Harmful Intervention.

Fig. 3.04. Image reference of the performance frameworks that were created for the evaluation.
The emphasis on including and describing these artifacts in detail comes from the need of establishing: a) objective, comparable metrics of performance; b) a methodology that may be easily replicable if the experiment is to be re-made in the future; c)research design that works as intermediate knowledge for future works, independent of PreViS.

3.6. Simulation Participants.

Participant Description:

In total, 6 participants were recruited. Control and intervention group had 3 participants each. Other relevant information about these participants is as follows:

- All participants were within the first year of their paramedic training.
- None of them had conducted a NIHSS examination before.
- All members in the intervention group had had a PreViS training session within the last month. Said sessions were 1 hour long.
- All participants were between 19 and 20 years old.
- The gender distribution was 5 females and 1 male.

.

Recruitment: All participants were recruited as volunteers. No incentives were offered. The sampling was one of convenience, although there was a purposeful selection process where paramedics were picked based on their professional experience level as well as their level of experience with PreViS.

Participant Management:

Participants were briefed on the general intention of the experiment. They were asserted that it was the system that was being evaluated, instead of their performance.

It was also made clear that their feedback and any other given data would be confidential and handled under GDPR protocols. An NSD agreement was provided for each of the participants to review and agree to. Efforts were made to placate possible reservations participants may have regarding giving honest feedback in fear of professional retributions.

3.7. Post-Simulation Evaluations.

After the observations were chronicled, the next step was to process the findings through an Usability Inspection method. **This was chosen as a pathway to answer the sub-questions in Q2.**

Q2: Is the Usability of the system acceptable for its use on prehospital emergencies? **Q2.1. Do users perceive the system to be easy to use?** **Q2.2. Is the use of the device causing an extra load (cognitive or otherwise) on the user.** ● Does the setup of the device respond to the user's mental model? Are there consistent user errors? **Q2.3. Is the HCI intuitive? (voice-command interaction, navigation etc).** ● Are the touchpoints of the interaction intuitive? ● Does the language used in the system (voice commands, sections labeling), match real world conventions? ● Is the system's navigation clear and accessible to the average user? **Q2.4. Is the interface prone to user errors?** ● Does the interface violate usability heuristics? **Q2.5. Is the system prone to technical errors?** ● Are there consistent malfunctions in the system?

As the first step to approach these questions, the researcher chose a Usability Inspection Framework.

3.7.1. Usability Inspection: Evaluation Framework.

For the usability inspection, 3 evaluation processes were implemented: A Heuristic Evaluation, a GUI analisis with emphasis in navigation, and the previously mentioned SUS Posttest.

3.7.1.1. Heuristic Evaluation.

A heuristic evaluation is a method which involves the comparison of an interface to a framework of pre-established design principles (Wheeler Atkinson et al., 2007). For a full definition on this matter, and the heuristic frameworks used in this research —for example, W. Atkinson's et. al. Multiple Heuristic Evaluation (MHET)—, please refer to the appendix.

3.7.1.2. GUI Analisis.

Another part of the inspection was dedicated to a heuristic evaluation of the GUI and the navigation of the system. The same aforementioned principles were applied. For the navigation inquiry, an Instone Test[12] was performed. Principles and recommendations from Susan Farrel's ''Navigation: You Are Here" (Farrel, 2015) were employed as the standard to evaluate the existing system.

For ease of reading, the findings of the heuristic evaluation and GUI analysis were portrayed in the following Evaluation Framework:

Issue Description: *Where the details of the issue are chronicled.*
Heuristic Violations: *A way to frame why these issues are notable in the context of usability, highlighting which rules they are violating.*
MoScoW Priority Level: *Based on the Moscow Methodology (see definition on the Appendix), it is a way to indicate which issues are the most critical or damaging for the status of the usability.*

3.7.1.3. System Usability Scale Posttest Questionnaire.

Lastly, as previously mentioned, in order to understand the self-reported perception of usability, a posttest, paper based System Usability Scale questionnaire (SUS) was given to the participants after the simulations to be answered anonymously, and in private. Aside from the typical questions of this framework, one blank field was added with: "Any other comment about your experience?" in order to capture any other feelings the participants wished to express. Definitions and further detail on the implementation of this tool can be found on the appendix.

The following section will chronicle the **results** found through the application of the Methodology.

---

[12] A definition of this test can be found on the appendix.

4. Results

4.1. Simulation results

The following section describes the events observed during the roleplay sessions. First, the observations from the Control Group will be summarized. This will be followed by the observations from the intervention group. For further reference, a full 30 page simulation-by-simulation report where the Performance Evaluation Frameworks depicted in the previous chapter is used, is included on the appendix.

4.1.1. Control Simulation Results.

4.1.1.1. Overview of Control Simulation Results.

The following is an overview of the results chronicled in the Evaluation Framework for the Control Simulations. Further detail from these results can be found in the subsequent sections.

Could NIHSS be practiced correctly?
No. In all 3 control simulations there were mistakes in the way the examination was practiced, either in the way the PMs implemented the movements and/or stimuli on the patient, and/or by missing and/or misreporting findings in the patient's pathology to the RHP.
Were there harmful Interventions? If so, how harmful were they?
Yes. As expressed above, the 3 simulations failed, in one way or another, in giving the correct information to the RHP. The consequences of this, within the scenario of a patient with a posterior stroke, might result in the patient not going to the hospital and getting the MRI they need, and being sent to a regular health center where they are to wait for a new assessment

instead. In this scenario, this delay can cost patients to experience possibly irreparable brain damage. It is because of these reasons that it can be said that there were harmful interventions and that the 'harm' to the patient could have been severe.
Was the overall examination efficient?
As it will be discussed in the following sections, there were efficiency issues in the ways the PMs handled the mobile phone-patient-RHP interaction (see ‘Observations on the Mobile Call + Speaker-Off configuration’). However, the more consequential errors came in the form of inefficiency in conducting the examination and then communicating the wrong information —An act that could have dire consequences for the patient's health—. Because of this, it is appropriate to say that the examinations were not at all efficient.

4.1.1.2. Observations on Control Simulations.

Fig. 4.01. Image of the Simulation showing Speaker-Off Configuration.

Observations on the Mobile Call + Speaker-Off configuration.

a) Speaker-Off mode = Participants out of the Loop:

While using a mobile phone to conduct the examination, one very important decision which sets the tone of the interaction is whether the Principal PM decides to put the telephone on speaker or not. Choosing not to use a speaker makes the PM the only person with all the information. All other participants wait in silence without being able to hear what the RHP is saying. It is not hard to imagine that this loaded silence and waiting period could generate feelings of anxiety or alienation in the patient. This setup also leaves the PM partner out of the loop on what is happening and possibly cutting on preparation or observation support.

b) Repetition in dialogue:

Because of the aforementioned configuration, the communication process can be somewhat redundant. This sequence works as follows:

(1)The neurologist asks a question.

(2) The PM repeats a question.

(3) The patient does the action.

(4) The PM reports the action back to the RHP.

For example:

(1) RHP: (on the phone): *"Can she do (x)?"*

(2) PM to the patient: *"Can you do (x) ?"*

(3) Patient: [Does the action].

(4) PM to Doctor: "*Yes she can".*

Having to repeat each question from the RHP to the patient and vice-versa can be an arguably unnecessary, time-consuming step in the examination process, which can make the interaction cost of the overall task to 'balloon up'. As will be explored in the following simulations, the inclusion of a speaker in the interaction reduces the need for this repetition.

c) One-handed Examinations:

Because the PM is holding the phone, they are conducting the examinations one-handed. This clearly limits the range of motion the paramedic can make, and may endanger the integrity of other, more complex examinations which need to be performed with both hands (such as Visual Field Testing Step #3 in the NIHSS Scale). While this is not an insurmountable challenge, it is less than ideal for paramedical examinations.

Observations on the Mobile Call + Speaker-On configuration.

a) Other Participants in the Loop and Less repetition.

The 'Speaker On' / hands free combination seems to make the operation work a lot more swiftly than the previous case. During these simulations the volume was high enough so that the patient could hear what the doctor was saying, and could interact directly with the doctor on the call. The sequence of interaction was as follows.

(1)The RHP asks a question,

(2) the patient answers the question directly.

Example:

(1) *"How old are you"?*

(2) "*53".*

(3) [the RHP continues with the next question]..

or

(1)The RHP asks a question,

(2) The patient does the action;

(3) the PM reports the action back to the RHP.

Example:

(1) RHP: (on speaker): *"can you close both of your fists"*?

(2) [Patient does so]

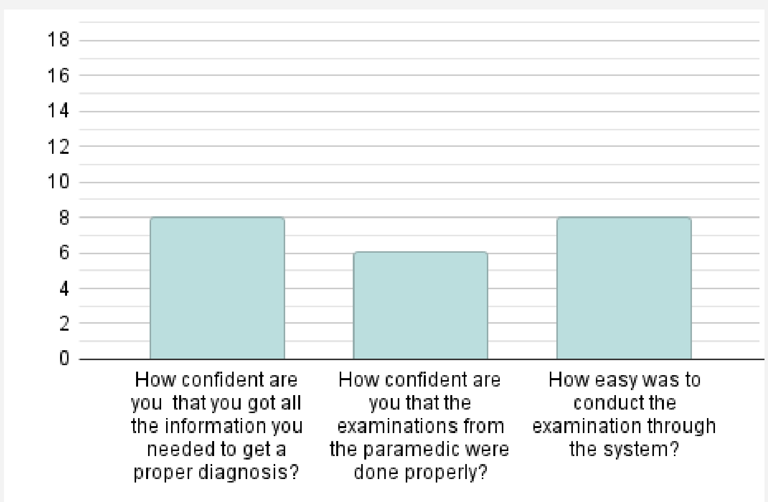
(3) PM: ***"****Yes*". [communicating to the RHP they did the action successfully].

b) Two-handed Examination.

Having the phone set on a surface allowed the PM to conduct the examinations with both hands instead of one, allowing for the correct implementation of these.

4.1.1.3. Feedback from RHP (Control Group).

The RHP scored all the feedback surrounding the control simulations on the lower half of the scale. This communicates that their confidence levels regarding the accuracy and ease of the examination were fairly low. Two out of three comments left by the RHP for these simulations talk about having to rely fully on the 'trust' that the PM made the examinations and observations in the correct and accurate manner.



	How confident are you that you got all the information you needed to get a proper diagnosis?	How confident are you that the examinations from the paramedic were done properly?	How easy was to conduct the examination through the system?
Control Group	8	6	8

Table. 4.01. Feedback from RHP (Control Group).

4.1.1.4. The Issue with the Status Quo.

Despite the improvements seen from the ‘Speaker-Off’ to the ‘Speaker-On’ mode simulations, the most crucial issue remains, which is that i**n the control group situations (i.e. the representation of the status quo of how emergency calls are handled in Norway), the integrity of the diagnosis depends on whether the PM is conducting this examination (and relaying the observed information) with accuracy —something they may or may not be trained enough to do—, so that the RHP can make an educated decision about the treatment of the patient.**

Each one of the control simulations had more than one examination error which compromised the possibility of carrying out NIHSS correctly. Said errors ranged from:

- Missing signs the patient was displaying: —for example, failing to recognize NIHSS symptoms (occuring in control simulation 1, 2 and 3) —

- Conducting the examination inadequately —for example, providing the stimuli too fast, or without achieving the required postures (occuring in simulation 1, 2 and 3)—

- Relaying erroneous information —for example, reporting a symptom happened on the left when it actually happened on the right (occuring in control simulation 2 and 3)—

This was a somewhat expected outcome because of the very low levels of experience that the participants had. However, it did illustrate how the current system of conducting examinations through the phone can be highly vulnerable to errors which may lead to patient harm.

The lack of visual component during the examination puts the entire onus on the PM to perform at 100% capacity and be able to:

> i) Perform the examinations correctly (e.g. performing more complex examinations they might not be familiar with, as it was the case with Step #3 and #8 of NIHSS: Visual Field Testing / Sensory, respectively).

As well as:

ii) Being able to notice the right information (e.g. noticing that one of the eyes did not respond to the stimuli).

iii)) Get a detailed enough explanation from the RHP on how to do the examination, and interpret the instructions in the way the RHP intended.

iv) Notice any other sign or symptom that might be relevant to the patient's condition which falls out of their professional scope.

The bottom-line is that the RHP is getting the information exclusively through second-hand dialogue, thus they have to trust that the information that is being relayed to them is correct, and that consequently they are making an informed decision. As recorded during observations, this is a fragile assumption.

In the following page, the **Intervention Simulation Results** can be found.

4.1.2. Intervention Group Simulation Results.

4.1.2.1. Overview of Intervention Simulation Results.

The following is an overview of the results chronicled in the Evaluation Framework for the Intervention Simulations. Further detail from these results can be found in the subsequent sections.

Fig. 4.02 Image from the intervention simulation.

4.1.2.1.1. Efficiency of Examination Evaluation.

Could NIHSS be practiced correctly?
Yes. All the stages of the examination were done correctly, and all patient symptoms were caught on camera and recognized by the RHP.
Were there harmful interventions? If so, how harmful were they

No harmful interventions were detected.
Was the overall examination efficient?
In all 3 of the simulations the following conditions were achieved: ● All the signs from the patient's condition were easily seen on camera by the RHP. ● NIHSS was performed successfully, and there were no harmful interventions. ● Any missteps from the PM were recognized and easily corrected (see image below). ● There was positive feedback from the RHP. In 2 out of 3 simulations the examination was done in a fluid manner, and there were no signs of the PM struggling with the headset. However, one of the simulations was tainted with technical glitches which clearly disrupted the rhythm of the examination (see 'Voice Command Failures' section). It is because of this that in this case the results to whether 'the overall examination was efficient' are inconclusive, and further data is needed to make a stronger assertion on the matter.

4.1.2.1.2. Headset Interactions: Technical Review.

Were there any instances of user error?
Yes, two main errors appeared in all 3 PreViS simulations.

First, during several instances the PM did not notice that the image they were transmitting was not aligned with their POV. The PM had to be reminded by the RHP that they had to make sure they were framing the relevant parts of the patient on camera. This points towards a possible mismatch between the user's mental model and the system's output. The issue will be discussed at length in the section 'The Role of the Cameraperson'. Second, while adjusting the microdisplay, the PMs inadvertently placed the mobile 'boom arm' in a way where it covered a small portion of the camera. Because of this, there was a small amount of occlusion on the video footage the RHP was receiving. This issue will be further analyzed in subsequent sections (see 'Occlusions on the Microdisplay').
Were there any delays or problems starting connection?
No.
Were there any instances of voice-command failures?
Yes. In 2 out of 3 simulations the voice command failures were minor, happening only once or twice. However, in one of the simulations, the voice command glitches became very prevalent and disruptive: Out of 14 commands recognized by the system only 8 of them were real/intentional commands spoken by the user. It is unclear what triggered these random activations of the zoom and light command. It's important to note that even though the examination was successful in that simulation, these occurrences were definitely very disruptive to the process.
Were there disturbances in Audio (Lag, interruptions, problems with volume, etc).
No instances recorded.
Were there notable fidgeting and readjustment instances with the headset?

The PMs were not seen struggling to put the headset on. All instances where the PM placed their hands on the device, it was to make minor adjustments which were resolved in a matter of a few seconds. Their observed body language did not indicate large gestures of discomfort. The only pervasive instance of adjustment happened in the beginning, when users participants were adjusting themselves to the mental model of what the camera was broadcasting, and, during the simulation, when prompted to adjust the framing of the patient (see 'The Role of the Cameraperson' section).

4.1.2.2. Observations on Intervention Simulation.

4.1.2.2.1. The Speaker Component

Using PreViS with the 'Speaker-On' option brings all the benefits of the Phone-Speaker-On configuration explored on the control group observations. Firstly, the PM can conduct the examinations hands-free; secondly, the RHP can interact directly with the PM and patient, 'looping-in' all participants, including the secondary PM.

4.1.2.2.2. The Visual Component

The effects of the visual element became relevant early on every examination. For example, during the first simulation the RHP was able to recognize the sign in step #2 Best Gaze (the moment when the RHP captures the eye movement the 'patient' was portraying can be clearly seen on the internal recording of PreViS). Second, the RHP was able to notice that the PM was conducting step #3 of NIHSS incorrectly. He/she stopped the examination, explained to the PM how to do it correctly, and then resumed, this time conducting the correct implementation. It is an example on how the concept of PreVis could compensate for the lack of experience and/or of errors / omissions committed in a prehospital environment.

Fig. 4.03. During all intervention simulations the RHP took the lead guiding the process. In the images above it can be seen a moment when the RHP is instructing the PM on how to make the examination. The PM moves their finger to the sides of the patient's face. This is an erroneous move, as the PM's finger should stay still and it is the patient who needs to move their head following the finger with their eyes. The RHP notices the error, corrects the PM and instructs the patient directly to move their head instead.

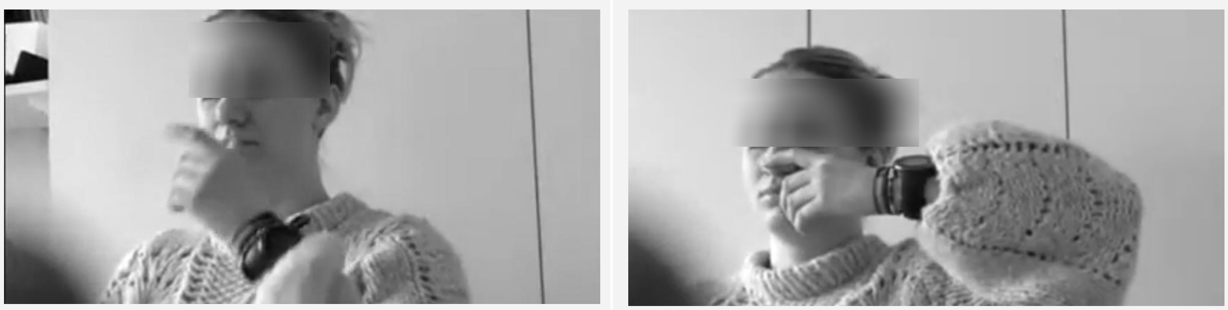


Fig. 4.04 The images above show the moment in the simulation when the RHP noticed the patient was not holding the right posture for the examination. He then corrected the patient themselves until the correct position was performed.

4.1.2.2.3. The Role of the Cameraperson: "Can you move the camera a little bit?

Because the video the RHP is receiving is tied to the user's head-movements, the PM must adapt to this atypical head-based cameraperson role. This became obvious everytime the RHP asked the PM to adjust the camera so they could see the target of their inquiry: the interaction of 'moving the camera a little bit to the [direction]' seems to be an intrinsic part of the PreVis-led examination process.

Being that the camera is mounted to one side of the headset and has lesser range and dynamics than a pair of human eyes, there is clearly a gulf between what the PM is seeing in real life, versus what they are actually capturing with the camera. However, the illusion/ assumption/ mental model that the headset is transmitting exactly the PM's point of view is strong enough so that the PM might often 'forget' (while busy with the examination), that this is not really the case, and have to be reminded by the RHP to readjust their perspective to fit the alignment of the camera[13]. This interaction goes as follows: During the examination, the RHP would stop and ask the PM to adjust the camera. The PM would check the live feed on their microdisplay and make the adjustment by moving themselves or the camera. Once the desired ‘frame’ or ‘shot‘ was reached, the RHP would typically give approval and continue the examination.

Fig. 4.05 These images capture an instance of the interaction where the RHP asked the PM to adjust the frame. In this case the RHP needed to see the patient try to smile to check for NIHSS #4: Facial Palsy. The interaction went as follows:

RHP: [to the patient]: “Can you smile?”

RHP: [to the PM]:“Can you adjust the camera a little bit so I can see the face?”

PM : [Adjusts the camera].

RHP: “That’s good”.

This is a relevant interaction to observe, as it is a new specific task added to the PM's workload. The second simulation made this point even more evident. While observing this PM's reactions / body language when adjusting the headset’s camera to comply with fitting the patient on the frame (tense, strained gestures; closing one eye to focus on the camera), one could see there was a relevant amount of effort dedicated to fulfill the RHP requests. This may indicate that there is a cognitive load (and/or perhaps a learning curve) towards understanding how to perform the typical (already complex) tasks of an emergency situation, while keeping the patient on camera, so that the RHP can see what they need to see.

[13] These reminders may or may not be a neutral or positive interaction in real life. They might even be stressful because of the urgency of the situation, or because of the power dynamics and or tone of the conversation between the PM and the RHP.

The PM closing one eye or concentrating both of them on the microdisplay, making sure the framing was correct, not only requires some physical / cognitive effort, but can also seemingly create some disconnect with the 'patient': **The person playing this role mentioned during interviews that he/she felt this PM was sometimes wholly concentrated on the headset instead of 'them'. They expressed that in their professional opinion, this may make some patients feel alienated from the process**.

Working on lowering this cognitive load, may overall benefit all the participants involved (in both performance, ease of use and possible adoption), and make the process more seamless.

Following this line of thought, the participant in the third simulation was a lot more adept at framing the shot than the previous two as he/she only needed one prompt in the beginning of the simulation to understand the relevance of transmitting 'the whole picture' to the RHP. Thus, he/she was always aware of the framing, and would move, change postures and position themselves dynamically according to what the part of the examination required.
In comparison to the other recordings, in this simulation one can see the relevant parts and appendages of the patient fully on the shot, according to each part of the examination. Despite the attention to this detail, the PM also managed to maintain a good connection with the patient. As it was explained by the person playing this role, this PM's bedside manner was strong enough in the beginning to establish a good connection, consistent eye-contact and there was a concise explanation of the steps to follow (i.e. the introduction of the headset, and process of the examination with the RHP) to make them feel reassured throughout.

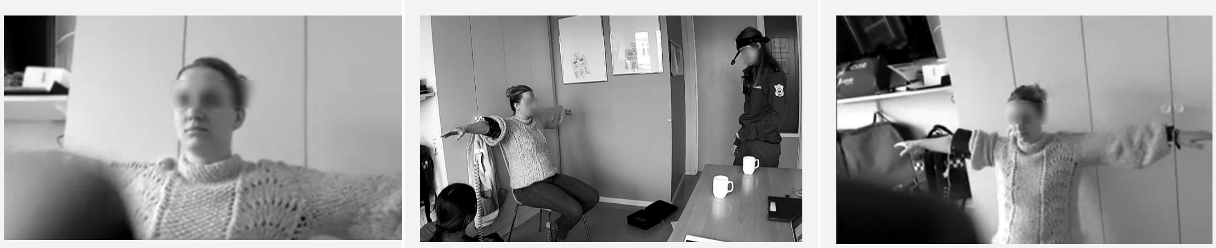


Fig. 4.06. Image of the third intervention participant, repositioning themselves without prompt in order to frame the patient's full span of arms for the RHP so they could evaluate NIHSS #8 limb ataxia.

The fact that this PM was able to strike a balance between 'being a good camera person' for the RHP and still making the patient 'feel seen' and not alienated from the process, may indicate that there is value in exploring and establishing best practices for the PMs to achieve both goals. In the case of the dynamic framing (changing postures unprompted in order to fit the relevant parts of the patient in frame), creating enough practice / muscle memory where the task may eventually have no significant/competing load on the overall process. In the case of bedside manner, implementing techniques on how to strike the right verbiage and overall interaction to improve the patient's experience throughout.

4.1.2.2.4 Voice Command Failures.

Different voice command failures were identified during the simulations. A few of these came in the form of voice commands uttered by the user which were not acknowledged by the system on ‘the first try’. However, the most prominent issue that surfaced during the simulation was that of the voice command failures. These consisted of random command activations on the headset without any prompt from the user. This malfunction was very prominent during the second PreViS simulation. In total, only 8 out of 15 commands actioned by the headset during this simulation were actually intentionally triggered by the user. The remaining 7 commands were randomly activated at different times of the examination. The majority of these commands were 'Zoom', —which made the image captured by the camera zoom in and out— and the other command was 'Light', which turned on the integrated flashlight of the headset[14].

Fig. 4.07. Images from the feed the RHP was receiving showing the moment the random ‘Zoom’ activations started to happen.

[14]All the roleplays were conducted with the same headset. It is unknown what made this roleplay more susceptible to the command-related technical glitches experienced here.

Fig. 4.08. Image from the simulation showing the PM struggling with the Headset's glitches. The zoom command going in and out randomly from the videofeed made framing the patient with the camera very challenging for the user, making them having to stop the examination altogether to adjust the view on the microdisplay.

It is not hard to imagine how disruptive these phantom activations were for the participants. At the beginning of the aforementioned simulation, the PM had voiced the fact that they were very nervous. These random activations (zoom, lights) certainly did not contribute to reassuring the PM.

The random zooming in and out of the image prompted by these random activations can be seen on the feed that the RHP was receiving. The fact that these could happen without the PM noticing could possibly generate a need for them to feel they have to constantly check the microdisplay to make sure everything is well framed for the RHP. This action can add an extra layer on the PM's workload, and or interrupt the connection or feeling of 'attention' between patient and PM. Furthermore, having the flashlight turn on and off randomly may possibly deter confidence from the work of the PM.

For the RHP who may not be familiar with the glitch of the headset, the random zooms on their video feed may prove to be a distraction, and may overall sour the tone of the interaction if they think the PM is the one responsible for unwanted changes in focus.

4.1.2.2.5. Occlusions on the Microdisplay.

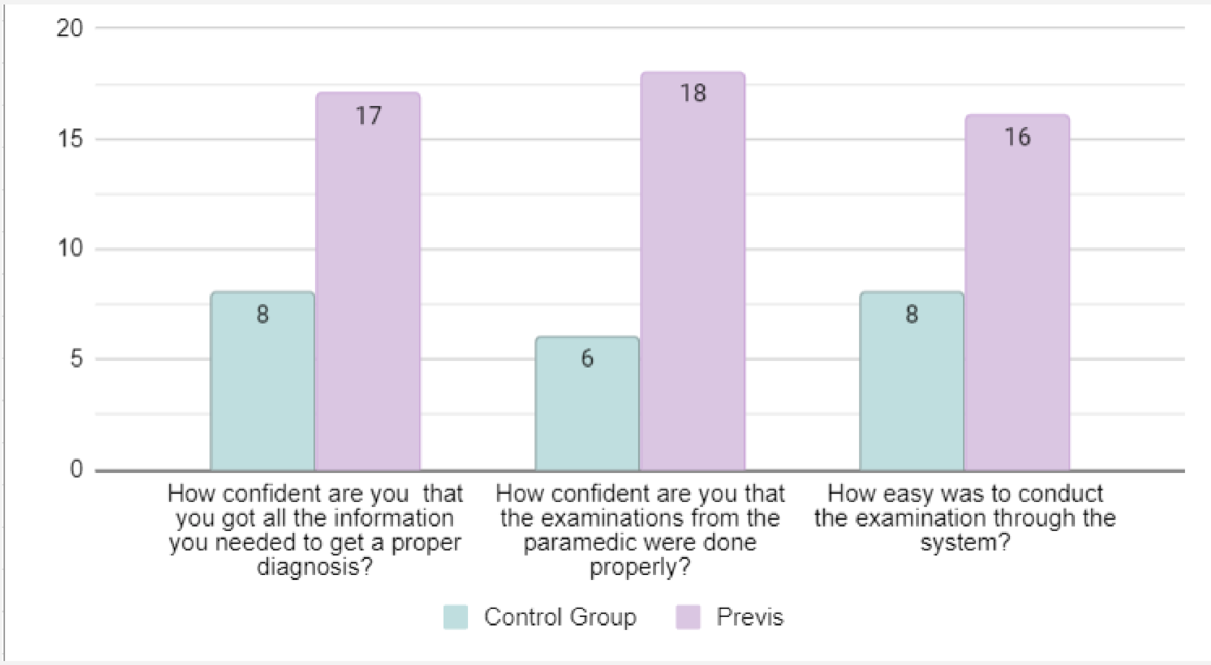
Because of the configuration of the device, the moving arm the microdisplay is fixed upon can be accidentally placed in front of the camera, thus covering a section of it. This can result in a small section of the video streamed to the RHP to be dark/hidden. Because of the configuration of the microdisplay, it might be hard for the PM to notice this if it hasn't been brought to their attention beforehand. This is a design flaw in the configuration of the headset, as there should be some physical constraint to prevent this error (consistent with the error-prevention heuristic). In the case of these particular simulations, it did not negatively affect the detection of signs from the RHP, however, if one of the main objectives of the device is to transmit visual information for diagnosis purposes, occlusions on the microdisplay are less than ideal.

Fig. 4.09. The image on the left shows the feed before the PM accidentally placed the 'boom arm' in front of the camera, causing the occlusion. The image on the right shows said obstruction on the bottom left corner. All three simulations suffered from this graphic disruption.

4.1.2.2.6. Overall RHP Feedback.

The RHP scored all the feedback regarding the PreVis simulations at the higher end of the scale, with the lowest median score being 16 points and the highest 18. This implies that the levels of confidence regarding the accruement of the necessary information to make an assessment, as well as the accuracy of the proxy examinations were on the high end of the scale. The same applies to their perception of how easy it was to conduct the examination through the system.

These scores contrast with the ones taken in the control simulation. Numbers-wise, within the PreViS simulation the confidence levels and ease of examination were doubled (and on one occasion tripled), in comparison to the control group. This implies that there might be an important difference in the RHP user experiences between the observed groups, having the PreViS group being evaluated much more positively.



	How confident are you that you got all the information you needed to get a proper diagnosis?	How confident are you that the examinations from the paramedic were done properly?	How easy was to conduct the examination through the system?
Control Group	8	6	8
Previs	17	18	16

Table 4.02. Feedback from RHP : Control vs. Simulation Group.

4.1.2.3. Results from SUS Posttest Questionnaire.

After the simulations, participants in the intervention group were given a paper-based SUS questionnaire. The data from these was then processed and calculated using the given method from the prescribed recommendations (see appendix for more details).The following are the results of said data.

	SUS Results		Median
1	**I think that I would like to use this system frequently.**		**4**
2	I found the system unnecessarily complex.		1
3	**I thought the system was easy to use.**		**4**
4	I think that I would need the support of a technical person to be able to use this system.		1
5	**I found the various functions in this system were well integrated**		**4**
6	I thought there was too much inconsistency in this system		2
7	**I would imagine that most people would learn to use this system very quickly.**		**4**
8	I found the system very awkward to use.		2
9	**I felt very confident using the system.**		**3**
10	I needed to learn a lot of things before I could get going with this system.		1
	Total SUS Score		**80**

Table 4.03. Results from SUS posttest questionnaire.

As previously mentioned, the results of the questionnaires must be taken with a grain of salt, as the sample size is so small, it is not wise to take this feedback as representative. For the purposes of this exercise, the following can be said:

The evaluation on the usability score is quite high, standing at 80 points. According to Bangor's et.al. assessments, this means the usability of the system is perceived as upwards of passable, in the 'better products' echelon. The feedback was stated at the extremes of the scale (4 for agreement 1 for disagreement). The highlighted prompts are written in a positive manner (i.e. I think *I would like to use this system frequently*). All the 'positive' prompts received the highest score for agreement, except for prompt number 9: *'I felt confident using the system'* which attained a 3.

For the negative prompts they all obtained the highest score or disagreement (1), except for two prompts: '*I thought there was too much inconsistency in the system*', and '*I found the system very awkward to use',* both obtaining a score of 2.

Unfortunately, none of the participants used the open question field which said “Any other comment about your experience?”.
More on these results will be examined in the discussion chapter.

4.2. Results from Usability Inspection: Heuristic Evaluation.

The following is a summary of the most relevant findings of the usability inspection. As explained on the Methodology chapter, these will be shown on the framework describing the following:

- Summary of the issue.
- Heuristic violated.
- Severity of error.

4.2.1.Usability Issues Found on Simulations.

4.2.1.1. Issue: The Camera-person’s Role and Toll.
In previous sections there has been discussion about the pressures of ‘being a good camera person’, and ‘keeping the patient in-focus’. To summarize, there seems to be a disconnect between what the interface promises to do, and what it actually does: Conceptually speaking, the implied premise of the headset is that it will transmit the user's POV to others. In reality, the headset is not broadcasting what the user sees, but a limited visual range a few inches away from the user’s temple. This disconnect between what the user expects and what actually happens has a negative impact on the experience, and puts the onus on the users to actively compensate for this flaw (the calls for readjustments of the camera, the intermittent checks to see if the relevant appendage was on the frame, etc). The gold standard for interfaces is to ‘become invisible’. That is to facilitate the primary task rather than make the user deal with the interface in order to facilitate it. Here, users are clearly dealing with the interface first in order to attend to the task second.

This, of course, has a cognitive toll, one that a busy paramedic may not have to spare. Observing the body language of some of the participants during the roleplay (strained faces, one eye closed, or both eyes focused on the microdisplay; long pauses for readjustments), one could theorize that said toll is not negligible. Further data is needed to assess whether this is the case, or even if this challenge is surmountable through spending more time using the device, i.e., if this is just a matter of a relatively attainable learning curve.
Summary of Problem: The camera broadcasts a different POV to what the paramedic is seeing. The paramedic has to be constantly aware of what they are broadcasting to the RHP, and /or make adjustments when the necessary visual is not 'on the frame'. **Heuristic Violations:** System-Real world match (MHET): Match between system and real world (Nielsen), and Metaphors (Tognazzini).
MoScoW Priority Level: **Should Have.** A paramedic attending an emergency call with a confounding patient (i.e. the scenario that PreViS was meant to cover), is already working at a high cognitive level, possibly in a stressful environment and having to make time sensitive decisions. A user in this context should not have to deal with the extra cognitive load of actively 'being a good camera-person'.

4.2.1.2. Issue: Ghost Commands and other glitches.
Random activations of voice commands were observed, particularly during one of the simulations. These were pervasive and disruptive to both PM and RHP. The source of these is unknown. One theory the developing team had is that the system (which is listening for commands in English) might mistakenly understand some Norwegian words as

commands. However, upon inspection of the video, some of the 'ghost commands' happened when no one was speaking.
In subsequent interviews it was revealed that other users have reported the same glitches on other PreViS/Realware headsets.
It is important to understand what is triggering these activation so they can be quelled in the future.
Another similar albeit less pervasive unwanted activation, happened when the system activated a command which was not uttered by the PM but by the RHP. There was only one incidence of this during the simulations, but this behavior was noticed twice in a previous training session as well.

Lastly, during the simulations, all commands were recognized in the first try except in 1 instance. On observations which predate the simulation session, this glitch was also observed, albeit only 3 times across 2 different users. This may be a negligible occurrence, but it is important to monitor for these failures in different contexts. It is particularly important to see if these glitches happen more or less often within different accents, and/or different English pronunciations. Contemporaneous experiences with voice command technology have shown that their tolerance to diverse accents can be a challenge (Coyle, 2018). On the version of PreViS that was submitted for evaluation for this paper, the list of voice commands has a mixture of english and norwegian terms (see this chapter's section on 'Inconsistency of Language.'). It is possible that users with less proficiency in spoken English or those who have a particular accent may be more vulnerable to these malfunctions.

Problem Summary: Voice commands present very noticeable malfunctions. These were shown to be very disruptive during examinations. The cause of said glitches is unknown.
Voice command is almost exclusively the only way the user can interact with the functions of the device. If this channel is defective, then the whole interaction suffers.

Heuristic Violations: Consistency. "*The Consistency heuristic emphasizes the importance of designing elements of an interface to provide standard and reliable terminology,*

actions[15] *and layouts*" (Wheeler Atkinson et al., 2007). Random glitches represent unexpected responses to the user's input. They taint the user experience by acting in unexpected, unreliable ways.

MoScoW Priority Level:
Must Have. This level of malfunctions is unacceptable for a device that has been released to users. Especially if it is marketed for use in medical emergency scenarios. When they happen, the malfunctions can be very noticeable, random and distracting. One could see this generating bad instances of user experience, and thus severely affecting initial user adoption.

4.2.1.3. Issue: Occlusion on the microdisplay.

Users may unknowingly and/or accidentally place the camera in a way where the mobile arm can block part of the image. During the simulations the observed occlusions were small, however, this incident happened in 2 out of 3 simulations, and said occlusions were part of the footage throughout the entire call. This is clearly a design flaw that should be corrected: For an equipment made to transmit important visual information, occlusions are less than ideal.

Heuristic Violations: *Error Management* (MHET), *error prevention,* and *help users recognize, diagnose and recover from errors* (Nielsen), et.al. (Wheeler Atkinson et al., 2007)

MoScoW Priority Level: Should Have.

[15] Emphasis is the author's own.

4.2.2.Usability Issues from GUI, IA and System.

The following findings do not derive directly from the simulations. They are part of a first-hand heuristic evaluation and analysis done on the device, or findings obtained while observing different paramedic groups interacting and/or training with the device.

4.2.2.1.Issue: Inconsistency of Language.
Interviews with experienced healthcare professionals in Norway indicate that the primary language of the user base of PreViS is Norwegian. Within the device most of the graphic elements in the screen, as well as the commands of the interface are in english, some others are in norwegian. There seems to be little logic on why and where one or the other are applied. For example: ● The *"HomePage*" interface text is in english (voice-commands in english). ● The "*Kompetanseavdeligen*" interface text is in Norwegian (voice-commands in English and Norwegian). ● The "*Call Space*" interface text is in Norwegian (voice-commands in English and norwegian). ● The "*Show Help*" menu text is in english (commands in english). ● The "*My Controls*" menu is in english (commands in english). This is not only a problem of inconsistency, but also an assumption that all the users will be comfortable reading, memorizing and speaking terms in a non-native language. As previously mentioned, there is also the assumption that their pronunciation will be ‘accurate enough’ for the system to recognize the command.
Heuristic Violations: Consistency. “*The Consistency heuristic emphasizes the importance of [...] standard and reliable terminology”* (Wheeler Atkinson et al., 2007).
MoScoW Priority Level: **Must Have.** The system must be consistent with the language it uses, and it shouldn’t

discriminate against users who are not proficient in a language that is not native to them, especially when voice-commands are the primary way users can interact with the device.

4.2.2.2. Issue: Non-intuitive, Inconsistent Voice-Command/Navigation Terminology.

As it will be explored in future sections, there are very little visual cues on the interface to remind the user of the voice-commands needed to navigate and interact with the interface. Some of the commands themselves can be very specific, non-naturalistic, bordering on jargon. Users will have to commit these specific terms to memory to be able to fulfill basic functions, such as going back in the navigation, or going back to the Homepage. If a user forgets what the term was they needed, they would have to guess, uttering command combinations with no feedback from the device[16] until they hit the 'magic word' that triggers the action. This is a scenario that even experienced users may encounter if they forget the name of a specific command.

For example, in the image below, only the terms highlighted in green will be recognized by the device. The terms in red will be ignored by it.

Fig. 4.10. Example of voice commands rejected (red) or accepted (green) by the system.

[16] There is no feedback when there is a command not recognized. While this is a violation of user feedback heuristics, this seems to come with the territory of a system that is always listening for commands, while the user might be conversing about other things.

Even within its own ‘technical language’ there are inconsistencies. The user can only navigate through different sections by saying: 'navigate home' or ‘navigate back’, however, 'navigate to help’, or ‘navigate up/down’ does not work. On a few instances during training sessions it was observed that the device will recognize the word ‘zoom’, but ignore the command ‘zoom-in’.

For in-page navigation purposes, the voice commands required for the user to trigger the actions are very specific. Only 'page down', "page up", "page left" and "page right" are accepted. More user-friendly variations such as the ones highlighted in red below will not be accepted by the system.

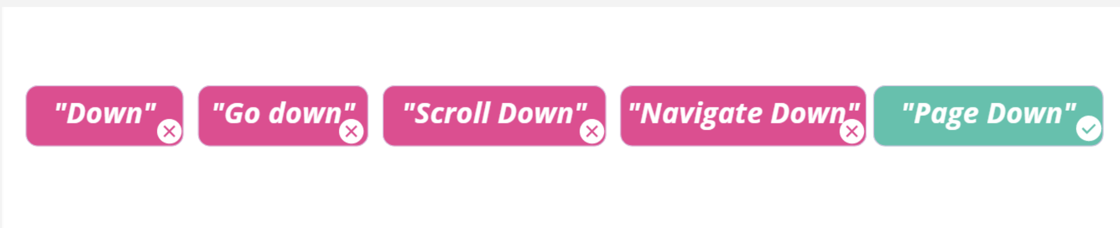


Fig. 4.11. Example of voice commands rejected (red) or accepted (green) by the system.

Changing settings on the menu, such as selecting volume level, and brightness levels, also requires a very specific prompt to select the desired option.

Fig. 4.12. Example of voice commands rejected (red) or accepted (green) by the system.

Even when uttering the right command : ‘Set level 5’, If there is a short pause (less than one second long) ‘Set level [Pause] 5’ will be ignored by the device.

Furthermore, within these configuration interfaces, if the user doesn't hit the 'magic words' in a short span of time, the screen will disappear, reverting back to the menu, making the user have to start over. This on itself is an example of inconsistency of feedback. Anecdotally speaking, as a first-hand account, the author can relate that if the user has not committed the terms to long-term memory, or has an instruction manual at hand[17], the 'guessing game' of trying to figure out the correct term to trigger the needed action can be a very frustrating experience.
Heuristic Violations: **Match between system and real world:** **"***The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon***. "** *(Nielsen, 2020).* "*The Consistency heuristic emphasizes the importance of [...] standard and reliable terminology"* (Wheeler Atkinson et al., 2007).
MoScoW Priority Level: **Should Have.** The system must be consistent with the language it uses, and it shouldn't discriminate against users who are not proficient in a language that is not native to them, especially when voice-commands are the primary way users can interact with the device.

After an analysis on the Information Architecture (IA) and navigational elements in the GUI, several flaws were detected. These will be summarized in the section below. A more detailed view, the full evaluation, including the findings of the Instone Test can be found on the appendix.

[17] This situation is more acute in the case of PreVis-specific commands, as (to the date of this publication) JodaPro has not made a manual available to their users.

4.2.2.3. Issue: GUI Navigational Flaws: Lack of Navigational Graphic Cues

Issue: The GUI does not offer enough support for the user to navigate intuitively, and fails at basic navigational strategies: "*Designers make navigation structures to help people move through websites, [...]. Navigation should not only show where you can go but also where you are now [...] It's important to convey enough context so that people can proceed immediately toward their goals*." (Farrel, 2015).

During the examination a noticeable lack of navigational elements was found. While conducting an Instone Test, it became very obvious that the user cannot rely on the GUI alone to understand where in the hierarchy each section is and/or what would be 'up' or 'down' from each level. There are no breadcrumbs, consistent section headings or other elements which can clue the users about this. Without knowing the necessary voice command by memory, the user would have to guess how to move to other levels in the hierarchy.

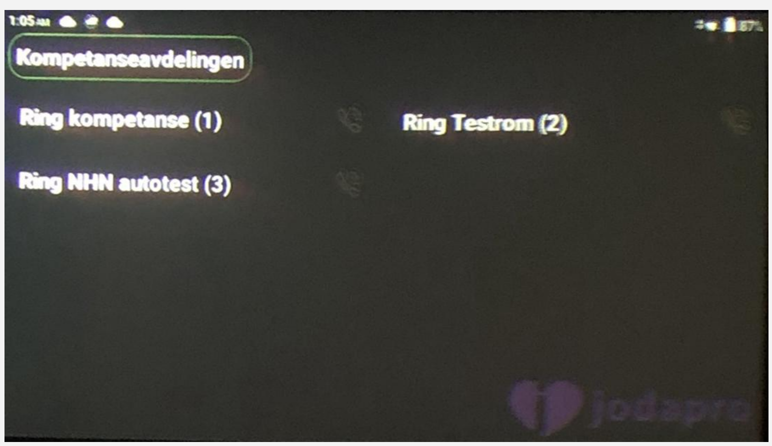


Fig. 4.13. Image of the Kompetensadelevigne UI. Notice the lack of navigational elements that would communicate to the users how to navigate from one section to another in this UI.

The GUI design strategy seems to be relying solely on 'Recall over Recognition' rather than the opposite. In terms of usability this is further aggravated by the fact that the interface is configured to be interacted by voice command only, so the user cannot rely on 'clicking around' to find their target or explore the IA by themselves. If they are unable to conjure 'the magic word' to coax the desired result from the system, the user will be stuck in the same section voicing commands indefinitely through trial and error.

This is why making always-visible navigational elements i.e. applying the 'recognition over recall' heuristic system wide is a high 'Should-Have' recommendation for the next iteration of the interface.

Heuristic Violations:
Software-User interaction / Cognition Facilitation, (Wheeler Atkinson et al., 2007) recognition over recall, (Nielsen, 2020).

MoScoW Priority Level:
Should Have. The lack of navigational cues severely infringes on the usability of the device.

4.2.2.4. Issue: Uninspected Accessibility of GUI

The main screen of the interface is a 0.33 inch microdisplay which rests at a range of a few centimeters from the user's eye. The screen itself is made of a magnifying microdisplay which amplifies the graphics on the screen. Because of this very specific configuration it was hard to apply traditional measures (WCAG, et.al.) to understand objectively how the UI was performing accessibility-wise, especially in terms of legibility, contrast[18] and font size.

From an anecdotal first-hand experience, the researcher believes readability can be compromised especially on graphic elements located towards the edge of the screen/microdisplay. These deformations are visible on the pictures taken from the interface, but they vary when the user tilts their head (in the same way one would see different magnifications when tilting one's head in front of a magnifying glass). Further observations / user data is needed to understand to what degree is this GUI accessible or not.

MoScoW Priority Level: Must Have.

[18] Some icons on the GUI are violating contrast accessibility recommendations.

It is important to know if the GUI fulfills accessibility requirements or not.

4.2.2.5. Issue: User entry point, and Lack of Supporting Documentation.

There is little to no intuitiveness present in the design of the device. It is not reasonable to assume that an inexperienced user can start using the device and complete basic tasks within a reasonable timeframe without any previous training, and/or access to an external manual. This is especially relevant taking into account the previous point about how specific basic voice-commands are.

This is exacerbated by the fact that to the date of this publication JodaPro has not made a manual available to their users[19].

The entry point to use this device currently depends entirely on the training that the Project Owner has provided first hand within physical training sessions. What is left after the training is finished is a single page with a list of steps on the basics of the system. This is a document that the PO created by himself on a word processor in order to compensate for the lack of an official manual. He prints them and places them on each case with the device before distributing them. The information about how to use PreViS then lives within this piece of paper, and the word of mouth left from what was captured by the listeners before the PO left the building.

Heuristic Violations:

- **Help and Documentation:** Interfaces should have methods and documentation which support the user to understand how to use the system independently. (Nielsen, 2020).
- **Defaults:** This refers to "*the importance of providing the user with default information as well as supporting easy adjustment of initial settings*".(Wheeler Atkinson et al., 2007).

MoScoW Priority Level: Must Have.

Since the device suffers from unintuitiveness, measures should be taken to support the user's

[19] The Realware headset has a manual, however the scope of this does not cover PreViS-specific functionality.

learning curve. The existence of this will help with proliferation of the system's use / adoption.

4.2.2.6. Issue: Lack of Error / User Feedback Gathering
There is no official way on which users of PreViS can flag bugs or glitches. Some instances of malfunctions (see "Voice-Command Glitches" section) have gotten reported through word of mouth back to the Project Owner. He has communicated these verbally to the developers. That is the extent of how user feedback is handled within the project. Because of this configuration, the user data is quite limited, and there is no way of knowing what relevant user experiences are falling through the cracks.
Heuristic Violations: This could obliquely be a violation of the Error Management heuristic (Nielsen, 2020), albeit in a more macro, conceptual perspective.
MoScoW Priority Level: **Must Have.** Users should have an easy way to flag bad user experiences. In the status quo, if there is a scenario of catastrophic system failure, it might take an undetermined amount of time for word of mouth to gather enough strength to reach the PO , and then the developing team. This disconnect might endanger user experience, reputation and consequent adoption of the system.

This closes the results section. In the following pages, a **discussion** of these results will be found.

5. Discussion.

5.1. Control vs Intervention Group Discussion.

5.1.1. Answering Q1: Comparison of Performance: Control Group vs PreViS Group.

In the following section, the answer of Q1 and their sub questions will be addressed.

Q1a) Can complex examinations be performed more accurately in comparison to the control group?

Q1b) Can the system help avoid patient harm?

The PreViS group performed better than the control group in every simulation in terms of the exercise of NIHSS. In the control group each simulation failed at performing the full examination correctly. In some instances the errors were in terms of performing the correct movements/stimuli to trigger the correct response on the patient (NIHSS steps #2 and #3 —best gaze and visual field testing— were particularly vulnerable to errors from the PM). Performing the examination one-handed (having one hand tied to holding the phone) was also found to sometimes compromise the integrity of said exams.

Other times, the errors were due to misreporting findings. These were, for example, when the PM failed to catch a sign, interpreted it as 'normal' and thus unintentionally gave misinformation to the RHP. Another example were instances where the PMs committed 'rookie' mistakes, such as reporting on signs based on 'their' right/left instead of the right/left of the patient (these happened with different participants in two separate control simulations). The consequences of these mistakes could have easily led to misdiagnosis and/or mild to severe patient harm.

5.1.1.1. The advantages of using speaker mode on the call.

Having the headset facilitating the speaker-mode on communication brings on several benefits to the examinations and patient interaction.

As explored in the control sessions, the use of the speaker has multiple benefits. Firstly, it allows the paramedic to conduct the examinations properly (i.e. using two hands). Secondly, It allows the patient and the partner-paramedic to be 'in the loop' and be active participants throughout the examination. This in itself has the following benefits:

- The patient can hear and respond directly to questions and instructions, cutting out the need for the repetition of the interloper (minimizing time, redundancies, and interaction costs). Furthermore, it can arguably make the patient feel less isolated from the experience.

- The Partner PM can be updated in real-time on what is happening in the examination/diagnosis, allowing possible contributions in observation support or pre-emptive equipment preparation.

5.1.1.2. The advantages of the visual component.

In the intervention group it was easy to see the immediate benefits that having a visual component can have on the examination process.

No patient harm was recorded in the PreViS simulations. There were no misdiagnoses because the examinations were conducted thoroughly —with any mistakes corrected in real-time—, and all the patient signs were caught directly on the camera and recognized by the RHP themselves. This helps bypass the risk of 'rookie mistakes'. For example where the PM would have misreported a sign being on the left when it is actually on the right, the RHP would have just seen the sign by themselves without the need of the flawed interpretation of the interloper).

In the previous chapter, it was explored what a paramedic must do in order to conduct all the tasks correctly in a phone-only examination. The following table compares how these challenges can be aided by the addition of the visual component.

What the PM needs to achieve when conducting the examinations through the phone.	How the Visual Component can aid this.
Perform the examinations correctly (regardless of whether they are familiar with them / have performed them before).	The RHP can guide the PM step by step on how to perform examinations they might be unfamiliar with. They can also see and correct in real-time if there are any deviations from the procedure.
Notice the right information (e.g. noticing that one of the eyes did not respond to the stimuli).	In the transmitted video feed, the RHP can see the symptoms in real-time, without the need to rely on a second-hand account.
Get a detailed enough verbal explanation from the RHP on how to do the examination; interpreting the instructions in the way the RHP intended.	The RHP can see if their instructions were clear enough / understood correctly by watching the performance of the PM. Deviations can be caught and corrected in real time.
Notice any other sign or symptom that might be relevant to the patient's condition which may have fallen out of their professional scope.	Giving the RHP a visual may provide them with more context or signs that might have fallen off the paramedic's 'radar'.

Table. 5.01. Table showing how the visual component can aid the vulnerabilities the PM can face.

Q1c) Can the symptomatology of the patient be effectively transmitted to the RHP?

To answer Q1c the following questions from the Evaluation Framework will be prompted.

- **Is the video good enough to capture examinations like NIHSS?**

  Following the reviews from the RHP and personal review of the footage captured by the headset, the answer is yes. All signs were easily captured on camera.

  The more detailed information related to eye coordination (best gaze and visual field testing) was easily recorded on camera.

Similarly, in step #4 Face Palsy, where the patient is instructed to frown, show their teeth and raise their eyebrows, all the aforementioned movements were easily distinguishable on camera without need of using the zoom function.

Fig. 5.01. Image from the headset feed showing the patient responding to the prompt of step #4 Face Palsy. All 3 prompts, frowning, raising eyebrows and smiling are easily recognizable on the transmitted image.

- **In general terms is the quality of video good enough to capture more detailed information?**

  For other examinations outside of NIHSS which require the review of more detailed signs such as pupil dilation and petechiae, there was not enough first person information/observations on the matter to come to a conclusion on one side or the other, as this fell outside of the original scope. Anecdotally, during the research interviews some RHPs mentioned that pupil dilation was able to be seen through the video, however this was not tested. It is recommended that trials analyzing this are run.

- **Were there disturbances in Audio (lag, interruptions, problems with volume, etc).**

  No audiovisual issues were ever reported or detected in the recorded video feed.

5.1.1.3. Summary: Answering Q1.

In terms of Q1:

Can the system help make examinations more accurate in comparison to the control group? ● **Q1a)** Can complex examinations be performed more accurately in comparison to the control group? ● **Q1b)** Can the system help avoid patient harm? ● **Q1c)** Can the symptomatology of the patient be effectively transmitted to the RHP?

All questions can be answered positively.
If the results observed during these simulations are found to be reliable, valid, generalizable and precise when applied to a larger and more diverse participant sample, it can be said that the PreViS system has a good proof of concept: whether through Realware hardware or equivalent, the implementation of this type of communication can possibly help minimize or even deter patient harm.

5.2. Usability Discussion: Answering Q2.

5.2.2. Answering Q2: Is the Usability of the system acceptable for its use on the PreHospital Emergencies?

5.2.2.1. Q2.1. Do users perceive the system to be easy to use?

The results from the implemented SUS seem to point out that the participants of the simulation perceived the usability of the system in a positive way (The full report on the interpretation of the

SUS results can be found on the Appendix). However, due to the low number of participants, the dataset is not strong enough to make assertions one way or the other. Furthermore, as discussed in previous chapters, the possibility of uncontrolled bias from the participants may also be affecting this data. Further research is needed to capture a more solid answer on this question.

5.2.2.2. Q2.2. Is the use of the device causing an extra load (cognitive or otherwise) on the user.

- **Does the setup of the device respond to the user's mental model? Are there consistent user errors?**

  Observations taken during the simulations seem to point to the existence of extra task loads on the users. For example, during the role-play sessions, paramedics were observed multiple times stopping the examination to adjust the camera in order to frame the patient correctly for the RHP. As mentioned in previous chapters, there's a disconnect between the user's mental model and the system's output; consequently the user is caught compensating for this (see 'the role of being a good cameraperson' section).

  These interruptions (obviously inexistent in the control group) suggest that there are new cognitive tasks playing a role in the interaction of PreViS; the interaction costs of such are not yet well known; however, the pauses of adjusting the camera through trial and error, and body language from the paramedics —frowning, closing one eye to focus on the microdisplay— suggest that the costs of these interactions may not be negligible. More data is needed to understand how big this toll is on the users, and whether it is a matter that lessens throughout the learning curve. Sourcing data such as a NASA TLX from a representative swath of participants may be a step in the right direction (see 'Future Work' section).

5.2.2.3. Q2.3. Are the touchpoints of the HCI intuitive (voice-command interaction, navigation wise)?

- **Are the touchpoints of the interaction intuitive?**

  The physical device is not noticeably hard to wear (users were never seen struggling with the fitting of such). The basic buttons can be easily found on the device, and their labels follow platform conventions. Observations never identified issues on this front.
  However, as mentioned in Q1.1. The microdisplay-camera interaction defies the natural mental model of the users, thus cannot be classified as intuitive.

  In terms of the GUI and voice interaction, during the Heuristic Evaluation section it was discussed that there is no ‘easy way in’ for first-time users, and it would be very hard for one of them to be able to fulfill basic functions without external instructions, or resorting to guessing the right commands to coax said functions from the system, as there are rarely graphic cues on the UI to help the user. It is clear the system is banking on a 'recall over recognition’ model (the opposite of the known heuristic), much to the hindrance of the overall usability.

- **Does the language used in the system (Voice commands, sections labeling, match real world conventions.**

  No. In the Heuristic Evaluation section there were several examples given on how the system uses jargon not only in the way it communicates with the user, but in the ways it expects the user to communicate with them: Being that voice commands are the only and very limited form of control/interaction users can exercise on the system, the user must become adept at speaking the device’s language in order to be able to access its basic functions. There is little to no tolerance for errors on the voice commands. To add to this, the language of the system is in of itself inconsistent, some terms being in english, others being in norwegian, This assumes that the users will: a) be fluent in english, when it is not the native language of the user base; and b) know when to use commands in english and when to use them in norwegian.

- **Is the system’s navigation clear and accessible to the average user?**

  No. The results from the Instone test and the GUI usability inspection reveal that there are very few visual cues to communicate to the user how to navigate through the system. This is another example of recall over recognition, where the user is expected to commit the information architecture of the site in their memory. During the simulations, the PMs were given a very basic and straightforward user journey, and were given 20 minutes before the task to familiarize themselves with these elements, thus the roleplay showed no noticeable struggles with the navigation. Because the Navigational UI structure is not intuitive nor explicit to the user (i.e. subpar to the recommended standards), it is reasonable to argue that an independent user without much training would be expected to struggle with the navigation if made to work independently with the device.

5.2.2.4. Q2.4. Is the interface prone to user errors? Does the interface violate usability heuristics?

Yes. In the Results chapter there were a number of heuristic violations highlighted which could easily lead to user errors. As mentioned previously, during the simulations, one of the most notable violations was regarding the disconnect between the user’s mental model and the functionality of the camera (metaphors / match between system and the real world). Furthermore, the aforementioned lack of user support and flaws on the GUI / navigation are primed real estate for user’s memory slips.

5.2.2.5. Q2.5. Is the system prone to technical errors?: Are there consistent malfunctions in the system?

Yes. While the audiovisual quality of the device serves its purpose well, the system’s malfunctions experienced during the simulations were very noticeable and distracting for all the primary users (PM, patient, RHP). These malfunctions were not exclusive to this simulation, as they had been noticed by the researcher during independent observations preceding this research. It is also an issue known to the PO and developing team. However, to the date of writing, the source of this error has not been identified or treated. Another less impactful technical error has to do with the ‘occlusions on the microdisplay’ finding, as they were found in all 3 intervention simulations.

In summary to answer the question: Is the usability of the system acceptable for its use on the PreHospital Emergencies?, the answer is no. It is the opinion of the researcher that this device is suffering from too many usability violations to be ready for launching. The prehospital environment is known for being hectic and unpredictable. The tools inserted to support these professionals must not add fuel to that fire. As mentioned in previous chapters, the best interfaces are invisible: i.e. those that allow the user to interact with the task itself rather than the intermediary tool. As seen on the simulations and usability Inspection, the current version of PreViS is not at that stage. Performing another reviewed iteration on the device's system would eliminate many of these flaws and potentially allow a more usable experience to the user.

6. Conclusion.

This research intended to understand the performance of the PreViS system in the prehospital environment. The scope of this analysis focused on answering the following questions:

Q1:Can the system help make examinations better in comparison to the control group?
Q2: Is the usability of the system acceptable for its use on Prehospital Emergencies?

For this purpose, simulations were held with the pursuit of observing and understanding the intricacies of inserting this new system and framework of interaction on what is already a highly complex activity.

To accompany these observations, posttest questionnaires were given to the participants (PM and RHP), in order to get some first-hand user feedback.

For a wider perspective on the user interaction quality within the system, an usability inspection was also conducted.

After these were implemented, the results from Q1 and Q2 were found to be mixed.

In the case of Q1, it became very clear early on that having a visual component while conducting complex examinations may bring an immediate upgrade in terms of accuracy in performance and interpretation of examinations. These factors consequently lower the possibility of patient harm.

If these results are found to be generalizable when tested through a larger, more diverse participant sampling, it may indicate a successful proof of concept for PreViS.

In the case of Q2, despite a very positive self-reported SUS score from the participants, the usability inspection applied to observations from the simulations, as well as the system's interface, navigation and/or GUI revealed several violations of heuristic and GUI best practices.

The overall outlook of this is that the system is not intuitive and will in probability prove to be challenging to interact with for first-time and/or infrequent users.

Perhaps the most alarming finding was that the current version of the device suffers from unpredictable technical malfunctions which were observed to disrupt the users performance. Said malfunctions were not unique to the simulations, as the researcher had noticed this behavior during observations preceding this body of work. Furthermore, interviews with the PO revealed that other users that have tried on the device had mentioned about experiencing these events as well. Despite the severity of this issue, to the time of this writing, the cause of the malfunction has yet to be identified, let alone amended. Understanding that this is an event which occurs randomly makes the prospect of it happening during a real-life emergency a dire yet possible scenario to envision.

From another point of view, observations during the simulation revealed that the interaction from the PM's perspective is not as simple (or intuitive) as it might have been expected.
Due to the fact that the PM is wearing a camera as a headset means that they are assuming the role of the camera-person, and they are solely responsible for appropriately framing the relevant body parts of the patient, all while conducting an examination on them. This is a new, added task to the PM's duties —duties which can sometimes be surrounded by very strenuous situations—.
The premise of the Realware Headset is that it will transmit the POV of the user. However, in reality, the limitations of the hardware do not allow for a 1:1 translation of this metaphor, and adjustments and care must be taken to fulfill a framing of the patient which can satisfy the dynamic requirements of the RHP; this effort must be maintained concurrently with the activities of attending, examining and treating the patient. There is currently not enough data to determine whether this is in fact a cognitive task whose load is heavy enough to weigh significantly on the paramedic. However, this new task has been identified, and it holds the potential to affect the cognitive load and/or performance of the paramedic. This means that it is a matter worth pursuing in further investigation in order to ensure that the tool is not ultimately hindering the user.

Q1 was answered positively while Q2 was answered negatively. These results are not incompatible. The concept of PreViS seems to provide more accuracy in examinations, consequently lessening the potential for patient harm. The observed issues mentioned in Q2 are all rooted in the hardware and software of the device, and run separate from the concept. In other words, the concept of PreViS can run independently from Realware's device and JodaPro's software. With the growing popularity of wearable devices, AR and body-mounted cameras, there might be value in researching —or designing— devices more apt for this activity, which can ultimately bring a more seamless —possibly less taxing— experience to the PM.

In the following section Limitations and Future Works will be discussed.

6.1. Limitations

Sampling and Biases

The strongest limitation of this research has to do with sampling. On one hand, the small number of participants blocked the possibility of handling a strong dataset, and achieving findings which are statistically significant and generalizable. In the same vein, recruiting further participants from different demographics/backgrounds would have provided a more representative image of how the system affected different user bases. For example, understanding how different age groups and levels of experience respond to the device (e.g.: Advanced users versus novices; frequent users vs. sporadic ones; Tech-savvy users versus less experienced ones).

On the other hand, a vast and diverse swatch of participants may have worked to ameliorate possible biases and professional pressures which may be affecting the data[20].

Due to the fact that the researcher had little control over the recruitment of the participants, and the PO was very involved during the process, one might argue that their presence could have somehow influenced the participants' behavior. This is not an atypical scenario in the world of device user-testing, as consultants often have to count on PO's for access to the user base, et. al. However, in order to ameliorate this, the researcher made efforts to triangulate

[20] During the early chapters of this document, it was described how the current role plays implemented by the Prehospital teams to pseudo-user-test PreViS lacked: a) structure, intentionality, replicability, systematic observation and analysis and documentation, and b) a way to ameliorate social dynamics which may affect honest user feedback, as well as a way to involve participants from outside of the silo in the process. While the matters of item 'a' were achieved, it can be argued that 'b' and 'c' were only partially accomplished.

methodologies, and create evaluation frameworks which would assist in creating objective observation and analysis.

The Silo.

One issue identified in the early stages of the research is that the PreViS project is being developed mostly in a silo of Prehospital Professionals. It was the original intent of the simulation that the patient and the RHP would be played by people outside of the paramedic Silo. In the case of the patient, this was achieved, as a nurse with multiple years of working first-hand with patients was able to fulfill said role. In the case of the RHP, the original intention was to recruit a Neurologist —I.e. The person who would be taking this phone call on a real 'suspicion of stroke' scenario—. However, due to logistical issues, the neurologist who was chosen for this role could not attend the roleplay session, and had to be replaced by a proxy: a veteran paramedic with experience implementing NIHSS.

This was unfortunate as the RHP is one of the main users of PreViS. A neurologist may have had different opinions and points of view which were imperceptible to other healthcare providers. In the future, involving other relevant members of the healthcare chain may bring to the table new unexpected perspectives and feedback.

The RHP POV.

The RHP participated in the simulation remotely. While their behavior was recorded in the footage of the video call, and their feedback was captured in posttest questionnaires, observing first hand the behavior of the RHP in their natural environment was not possible in this iteration of the research. Had there been a second researcher available, it would have been interesting to witness the experience from the RHP's POV, as it could have led to relevant findings about this side of the interaction.

6.2. Future Work

There are many paths where to expand on this research. In the following pages, some of these are highlighted.

Q3: The Wearability Question.

On earlier iterations of this research there was a Q3 which aimed to inspect the wearability of the device. For this, the researcher intended to use measures such as The Borg Rating of

Perceived Exertion (RPE) scale, (Williams, 2017), and/or the CRS (Comfort Rating Scales), (Cancela et al., 2014), in order to measure perceived comfort, exertion, etc. However, having only 3 participants to source quantitative feedback from would have not led to any statistically significant findings.

If these studies were to be re-made then an assessment of the wearability of the device might be useful, particularly, in understanding if there is any potential wearing the device may cause any sort of bodily harm (head, neck or eye strain etc.).

Q4: The Task Load Issue.

Similarly to the topic above, early iterations of this research intended to use the NASA TLX to understand the Task Load inflicted on the PMs by using PreViS. This approach suffered from the same ‘lack of participant syndrome’, and thus was discarded. On a theoretical new iteration of the simulations with a larger participant base, implementing a NASA TLX posttest questionnaire might prove useful in understanding this very important dimension of usability and human factors inherent in the system. Furthermore, an interesting study on this track would be to conduct a NASA TLX on a longitudinal study. This would allow researchers to analyze the shape of the learning curve in relation to the device's task load.

Observing other setups or configurations.

What happens if the partner PM is the one wearing the headset? Does this allow the primary PM to concentrate on the examination while the secondary PM concentrates on ‘being a good camera-person? Would the RHP get enough quality of video with this configuration? Would this choreography between the two PMs be easy to coordinate?

Other Interfaces.

The focus on this research was limited to the headset. It would be interesting to analyze the interface that the RHP is interacting with. How is said interface performing? Is it meeting the needs of the user? The same thought extends to the parts of the PreViS system which involve the cameras in the ambulance.

7. Bibliography.

Accorsi, T.A.D., Moreira, F.T., Pedrotti, C.H.S., Amicis, K.D., Correia, R.F.V., Morbeck, R.A., Medeiros, F.F., Souza, J.L. de and Cordioli, E. (2022). *Telemedicine diagnosis of acute respiratory tract infection patients is not inferior to face-to-face consultation: a randomized trial.* Einstein (São Paulo*),* 20. doi:10.31744/einstein_journal/2022ao6800.

Bangor, A., Kortum, P.T. and Miller, J.T. (2008) *An empirical evaluation of the system usability scale*, International Journal of Human-Computer Interaction, 24(6), pp. 574–594. Available at: https://doi.org/10.1080/10447310802205776.

Bergrath, S., Reich, A., Rossaint, R., Rörtgen, D., Gerber, J., Fischermann, H., Beckers, S.K., Brokmann, J.C., Schulz, J.B., Leber, C., Fitzner, C. and Skorning, M. (2012). *Feasibility of Prehospital Teleconsultation in Acute Stroke – A Pilot Study in Clinical Routine*. PLoS ONE, 7(5), p.e36796. doi:10.1371/journal.pone.0036796.

Cicero, M.X., Whitfill, T., Walsh, B., Diaz, M.C., Arteaga, G., Scherzer, D.J., Goldberg, S., Madhok, M., Bowen, A., Paesano, G., Redlener, M., Munjal, K., Kessler, D. and Auerbach, M. (2018). *60 Seconds to Survival: A Multisite Study of a Screen-based Simulation to Improve Prehospital Providers Disaster Triage Skills*. *AEM Education and Training,* [online] 2(2), pp.100–106. doi:10.1002/aet2.10080.

Courage, C., Baxter, K. and Caine, K. (2015). *Understanding your users : a practical guide to user research methods*. Amsterdam ; Boston: Elsevier, Morgan Kaufmann.

Coyle, A. (2018). *Speech Recognition: Affordances and Limitations with Accented Speech | CCTP 711 – Semiotics and Cognitive Technology*. [online] CCTP 711 – Semiotics and Cognitive Technology. Available at: https://blogs.commons.georgetown.edu/cctp-711-fall2018/2018/12/11/speech-recognition-afforda nces-and-limitations-with-accented-speech/ [Accessed 1 Dec. 2022].

Farrel, S. (2015). *Navigation: You Are Here*. [online] Nielsen Norman Group. Available at: https://www.nngroup.com/articles/navigation-you-are-here/ [Accessed 1 Dec. 2022].

Gibbons, S. (2021). *5 Prioritization Methods in UX Roadmapping*. [online] Nielsen Norman Group. Available at: https://www.nngroup.com/articles/prioritization-methods/ [Accessed 1 Dec. 2022].

Höök, K. and Löwgren, J., (2012). *Strong concepts: Intermediate-level knowledge in interaction design research*. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *19*(3), pp.1-18.

Instone.org. (2021). *Instone Test: Navigational Stress Test*. [online] Available at: http://instone.org/navstress [Accessed 1 Dec. 2022].

Ments, M.V., (1999). *The effective use of role-play: a handbook for teachers and teacher trainers-Revised Edition*. *Kegan Page Limited. London*.

Nestel, D. and Tierney, T. (2007). R*ole-play for medical students learning about communication: Guidelines for maximizing benefits*. *BMC Medical Education*, 7(1). doi:10.1186/1472-6920-7-3.

Nielsen, J. (1995). *Technology Transfer of Usability Inspection Methods (Jakob Nielsen)*. [online] Nielsen Norman Group. Available at: https://www.nngroup.com/articles/technology-transfer-of-heuristic-evaluation/ [Accessed 1 Dec. 2022].

Nielsen, J. (2020). *10 Heuristics for User Interface Design*. [online] Nielsen Norman Group. Available at: https://www.nngroup.com/articles/ten-usability-heuristics/ [Accessed 1 Dec. 2022].

Price, R., Bendall, J.C., Patterson, J.A. and Middleton, P.M., (2013). *What causes adverse events in prehospital care? A human-factors approach*. *Emergency Medicine Journal*, *30*(7), pp.583-588.

RealWear. (n.d.). *REALWEAR HMT-1®.* [online] Available at: https://www.realwear.com/hmt-1/ [Accessed 1 Dec. 2022].

Rivera, G., Sæbø, S., Van der Bos, L. (2021) *Roleplaying PreViS; Creating A Roleplay Framework for Improved User-Testing the PreViS System in Pre-Hospital Services.* NTNU.

Rörtgen, D., Bergrath, S., Rossaint, R., Beckers, S.K., Fischermann, H., Na, I.-S., Peters, D., Fitzner, C. and Skorning, M. (2013). *Comparison of physician staffed emergency teams with paramedic teams assisted by telemedicine – a randomized, controlled simulation study*. Resuscitation, 84(1), pp.85–92. doi:10.1016/j.resuscitation.2012.06.012.

Söderholm, H.M., Sonnenwald, D.H., Cairns, B., Manning, J.E., Welch, G.F. and Fuchs, H. (2007). *The potential impact of 3d telepresence technology on task performance in emergency trauma care*. Proceedings of the 2007 international ACM conference on Conference on supporting group work - GROUP '07. doi:10.1145/1316624.1316636.

Stickdorn, M., Hormess, M. E., Lawrence, A., & Schneider, J. (2018). *This Is Service Design Doing: Applying Service Design Thinking in the Real World* (1st ed.). O'Reilly Media.

Summers, A. and Willis, S. (2010). *Human factors within paramedic practice:the forgotten paradigm*. Journal of Paramedic Practice, [online] 2(9), pp.424–428. doi:10.12968/jpar.2010.2.9.78627.

Summers, A. and Willis, S., (2010). *Human factors within paramedic practice: the forgotten paradigm*. Journal of Paramedic Practice, *2*(9), pp.424-428.

Sykehuset Innlandet HF. (2021). *PreVis.mp4*. [online] Vimeo. Available at: https://vimeo.com/546394059/0d4efece26 [Accessed 11 Dec. 2022].

Valenzuela Espinoza, A., Van Hooff, R.-J., De Smedt, A., Moens, M., Yperzeele, L., Nieboer, K., Hubloue, I., de Keyser, J., Convents, A., Fernandez Tellez, H., Dupont, A., Putman, K. and Brouns, R. (2016). *Development and Pilot Testing of 24/7 In-Ambulance Telemedicine for Acute Stroke: Prehospital Stroke Study at the Universitair Ziekenhuis Brussel-Project*. Cerebrovascular Diseases, 42(1-2), pp.15–22. doi:10.1159/000444175.

Wheeler Atkinson, B.F., Bennett, T.O., Bahr, G.S. and Walwanis Nelson, M.M. (2007). *Development of a Multiple Heuristics Evaluation Table (MHET) to Support Software Development and Usability Analysis*. Lecture Notes in Computer Science, pp.563–572. doi:10.1007/978-3-540-73279-2_63.

Williams, N. (2017). *The Borg Rating of Perceived Exertion (RPE) scale*. Occupational Medicine, [online] 67(5), pp.404–405. doi:10.1093/occmed/kqx063.

(N.A) (n.d.). NIHSS. [online] Available at: https://www.nihstrokescale.org/.(Accessed: 1 November 2022).

(N.A) (n.d.) *System Usability Scale (SUS)* [online] Usability.gov. Available at: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html (Accessed: 1 November 2022).

Appendix.

1. Other specifications for Realware's Headset.

- 16 MP 4-axis optical image stabilization. Video output: Up to 1080p @30fps.
- 20º field-of-view. 1 meter fixed focus. 24-bit color LCD, 0.33 inch diagonal. Outdoor visible WVGA (854x480).
- 4 Digital microphones programmed for noise cancellation. The Device specification documentation reports "Accurate voice recognition even in 95 dBA of typical industrial noise".
- Internal 91dB speaker.
- Embedded LED flashlight.
- PPE accessible. Can be used with glasses and or regular eyewear.
- Reported shift-long (9-10 hours) lasting battery.
- Connectivity Specification: Wi-Fi: 802.11 a/b/g/n/ac – 2.4GHz and 5GHz. Bluetooth Low Energy 4.1.
- The PreViS setup includes a ready to use portable Modem.
- The device and accessories come in an easy to carry case.
- The device's headband is optional and it can be removed and easily cleaned.

---

2. Definitions.

2.1. Heuristic Evaluation:

A heuristic evaluation is a method which involves the comparison of an interface to a framework of pre-established design principles (Atkinson, T; Bennett, G. et. al. 2007).

There are many well-regarded and robust sets of heuristics. For this research the framework used was one of aggregation. Because there are many areas where multiple design guidelines overlap, B.F. Wheeler Atkinson et al.created a framework which consolidated what they considered to be the more robust principles into one structure:

"[...]there is a high degree of overlap among [heuristic] approaches. With this in mind, the authors integrated and updated these approaches in a single heuristic table – the Multiple Heuristics
Evaluation Table (MHET) to simplify access to a comprehensive set of heuristics
Concepts" (Atkinson, T; Bennett, G. et. al. 2007 p 585).

The MHET was the selected framework to which the observation of the simulation was evaluated against.

2.2. Instone Test.

Named after its creator, Keith Instone, this test is a method of measuring how easy it is for an user to understand the Information Architecture of a digital interface. It is composed of 10 questions which interrogate the graphic cues which can guide the user towards an easy navigation. The core of the questions can be summarized in the following 3 inquiries. These represent *"3 basic concerns users often have upon arriving at a page*" (Instone.org, 2021).

- Where am I?
- What's here?
- Where can I go?.

These inquiries represent a stress test on how intuitive the navigation of an interface is for an user.

2.3. System Usability Scale (SUS).

Developed by John Brooke in 1986, the SUS is a measuring tool created to capture the users' self-reported perception of the usability in a product (Bangor, A et.al 2008).

The implementation is fairly straightforward: For their scoring, participants are given the following 10 questions and asked to rate them with a range of 5 degrees in a spectrum from 'Strongly Agree' to 'Strongly Disagree'.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

For interpreting the scores the procedure is the following:

"The participant's scores for each question are converted to a new number, added together and then multiplied by 2.5 to convert the original scores of 0-40 to 0-100. Though the scores are 0-100, these are not percentages and should be considered only in terms of their percentile ranking". (Affairs, A., N.d).

Despite its simplicity, the SUS has been validated as an industry standard, as their results have been proven to be quite strong. Bangor A. et. al. conducted a study gathering 10 years worth of SUS evaluations across a wide breadth of product types. Their data analysis shows that the SUS can be quite a robust tool of usability evaluation. (Bangor, A et.al 2008).

2.4. MoScoW Methodology.

The MosCoW method is a framework of prioritization developed by Dai Clegg in 1994 typically used as a guideline to classify the requirements of a design and or development project. Their title makes reference to their classification criteria (Must Have, Should Have, Could Have, and Will Not Have), which Sara Gibbons (2021) describes as follows:

- "***Must have:*** *items that are vital to the product or project. Think of these as required for anything else to happen. If these items aren't delivered, there is no point in delivering the solution at all. Without them the product won't work, a law will be broken, or the project becomes useless.*

- ***Should have:*** *items that are important to the project or context, but not absolutely mandatory. These items support core functionality (that will be painful to leave out), but the project or product will still work without them.*

- ***Could have:*** *items that are not essential, but wanted and nice to have. They have a small impact if left out.*

- ***Will not have****: items that are not needed. They don't present enough value and can be deprioritized or dropped"* *(*Gibbons, S. 2021).

In the context of this research, this methodology will be used to provide a sense of severity or lack thereof on the findings of the research.

3. Performance Evaluation Frameworks.

After each simulation, the events were chronicled in an overview, and then interrogated by the questions of the framework. Said questions aimed to give enough background to answer Q1 and Q2 more thoroughly.

The emphasis on including and describing these artifacts in detail comes from the need of establishing: a) objective, comparable metrics of performance; b) a methodology that may be easily replicable if the experiment is to be re-made in the future; c)research design that works as intermediate knowledge for future works, independent of PreViS.

Titles in **bold** represent the questions that were asked of each simulation. The text in *italics* represents the reasoning behind this inquiry.

Items marked with a (**) were exclusive to the intervention simulations.

3.1. Performance Evaluation Framework: Efficiency of Examination.

Q: Was the Examination Efficient?
Why ask this?: The section named "Efficiency of examination' will interrogate the quality of the examination in both control and intervention groups.   ***Performance Definition:*** *Efficiency is a somewhat loaded term to introduce into the framework, as it is usually associated with time measurements.* ***For the purposes of this research, we are defining efficiency in a more holistic way****,* ***which can encompass the complexities of what performance is in this context.*** *Some of the questions that Efficiency is interrogating in this matter are:*   ● *Was NIHSS performed successfully?*   ● *Were there Harmful interventions?*   ● *Was there good feedback from the RHP regarding the examination process?*

- *Was there good feedback from the Patient regarding the examination process?*
- *Could all signs be transmitted and recognized by the RHP on camera? (**)*
- *Was the PM seen struggling in any way with the system/device? (**)*

If these questions are answered positively, then the performance can be qualified as efficient.

Q: Could NIHSS be practiced correctly?

Why ask this?: It is important to understand how complex examinations such as NIHSS are conducted when being guided by an RHP. This will allow an understanding on the status quo of the performance (control group), and ease the comparison process with the introduction of PreViS (intervention group).

Q: Were there harmful Interventions? If so, how harmful were they?

Why ask this?: The ultimate goal of PreViS is to reduce patient harm. It is important to chronicle incidents for comparison purposes (Control vs. Intervention).

Performance Definition: *For these measurements we are defining 'harmful interventions' by borrowing Söderholm definition: " [these are] treatments or procedures that are potentially harmful for the patient by increasing risks for complications or unwanted side-effects and by delaying the time it takes to the [patient's to reach a stable condition]" (Söderholm, M. 2013 pp. 3).*

In more straightforward terms, if the PM:

- *Conducts parts of the examination in an erroneous manner.*
- *Fails to notice signs the patient is portraying.*
- *Communicates wrong information to the RHP.*

These actions qualify as a Harmful Intervention.

Q: Overall, was the process efficient?
Why ask this?:   *This question allowed a more holistic look at the entire process. NIHSS may be practiced correctly, and there can be no harmful interventions, however, for example, if the process was marred with system malfunctions, if the user had to struggle to get the device going, and/or if the examination has significant delays, et.al. one would not have been able to say that the examination was efficient.*

3.2. Performance Evaluation: Headset Interactions: Technical Review.

Q: Were there any instances of user error?(**)
Why ask this?:   *Chronicling user errors is a good way to understand if the mental model the device/system communicates is effective enough, or if/where it is failing.*
Q: Were there any delays or problems starting connection?(**)
Why ask this?:   *Problems establishing a connection may significantly hinder the process of time-sensitive examinations.*
Q: Were there any instances of voice-command failures?(**)

Why ask this?: *During observations preceding the simulations, instances of voice-commands malfunctions were detected: (e.g. the system failing to recognize a command, or 'recognizing' one when no word was uttered). Chronicling these events in a controlled simulation will help understand their frequency and consequential effect in the context of an emergency scenario.*
Q: Were there disturbances in Audio (Lag, interruptions, problems with volume, etc)?(**)
Why ask this?: *The core function of the device is to send and receive audiovisual information within a time-sensitive context. Consequently, factors that hinder the quality of said information will affect the overall usability of the system.*
Q: Were there notable fidgeting and readjustment instances with the headset?(**)
Why ask this?: *Notable fidgeting and readjustment of the device may imply the fit of it (wearability) is somehow inadequate. This inadequacy may range from simple fit issues, perceptions of cumbersomeness, to actual discomfort. All of the aforementioned may carry a cognitive cost which might be subtracted from the attention which should be dedicated to the process of diagnosis/treatment, not to mention they may indicate bodily harm to the user, endangering overall usability.*

3.3. Performance Evaluation Framework: Feedback from RHP.

This section corresponds to a post-test questionnaire the RHP was given to answer after each simulation, in both control and intervention groups.

Some questions may be similar to those listed in the section preceding this. However, the purpose of these is to understand the experience from the specific RHP point of view.

Q: Did you experience any miscommunications?(**)

Why ask this? *Miscommunications may indicate either an audio problem, or a flaw/readjustment needed in the communication loop which may be intrinsic to the interaction in this unfamiliar device.*
Q: Was there ever a drop in quality of video or audio that got in the way of examinations. (**)
Why ask this? *This intended to monitor user feedback on the audiovisual quality the RHP was receiving*.
Q: Was the image of the patient in focus enough so that you could see what you needed from the patient?**
Why ask this? *This line of questioning was intended to understand if the image the RHP was receiving was clear enough for them to capture the details necessary to reach an informed diagnosis.*
Q: Any other comments about your experience? (**)
Why ask this? *A space for the participant to express opinions not encompassed by the questionnaire.*
Q: How confident are you that you got all the information you needed to get a diagnosis **Q: How confident are you that the examination from the paramedic was done properly?**
Why ask this? *These two questions were intended to capture any reservations the RHP may have had about the perceived level of accuracy of the examination.*

Q: How easy was it to conduct the examination through the phone/headset?

Why ask this?

The perception of ease of use is vital to usability. Including this on the control group experience will allow to chronicle the RHP experience in both circumstances.

Q: Any other comments about your experience?

Why ask this?

A space for the participant to express opinions not encompassed by the questionnaire.

---

4. Control Simulations.

4.1. Control Simulation 1

4.1.1. Overview of Simulation 1

The simulation began. When the phone call stage of the process started, the PM chose to take the call with the mobile to their ear, i.e. without the use of the speaker. Because of this decision, the only person that had instructions and information from the RHP throughout the whole process was the PM. The rest of the participants were waiting in silence unless prompted. There were long, silent pauses while the RHP explained to the PM how to do each part of the examination.

The PM conducted the examinations one-handed, as he/she was holding the phone to their ear for the duration of the call.

During the examination, the paramedic was prompted to do a 'visual field test' examination (Step #3 for NIHSS). The PM erroneously conducted the examination corresponding to 'best gaze' (Step #2 for NIHSS) again instead. Because of this mistake, the PM failed to recognize a symptom that the patient was presenting (introduced in the simulation to signal a posterior stroke), and thus communicated the wrong information to the RHP. Due to the lack of visual feedback, this mistake was unbeknownst to the RHP.

Similarly, in step #8 of NIHSS, the paramedic was meant to conduct the examination on both the upper body part, as well as the lower half. The PM only conducted the examination from the waist up, neglecting to test the lower limbs for signs of sensory loss. This neglect could have led to a critical misdiagnosis / a harmful intervention.

The rest of the examination was carried out normally, and the roleplay concluded.

4.1.2. Results

4.1.2.1. Performance of Examination

Could NIHSS be practiced correctly?

No. Both Step #2 (Best Gaze) and Step# 8 (Sensory) were done incorrectly. In the case of the former, this mistake prevented the PM from 'catching' the sign that the patient was presenting in the roleplay. The RHP on the other side of the line would have not known that this was the case and that thus the information communicated to him was incorrect.

Were there harmful Interventions? If so, how harmful were they

The quality of the examination was deficient, and detection of some signs was not made. This shows how the RHP could have been led to reaching a wrong diagnosis. Theoretically, the impact of this may vary from delays in time-sensitive treatment and/or severe consequences from a misdiagnosis.

In this case, failing to recognize signs of a posterior stroke might result in the patient not going to the hospital and getting the MRI they need, and being sent to a regular health center where they are to wait for a new assessment instead. In this scenario, this delay can cost patients to experience possibly irreparable brain damage.

It is because of these reasons that it can be said that there were harmful interventions and that the 'harm' to the patient could have been severe.

Was the overall examination efficient?

The PM's decision to hold the phone to their ear kept the other participants from being 'in the loop about what was happening, resulting in some otherwise unnecessary repetition. Furthermore, holding the phone in one hand limited their ability to conduct the examinations properly. However, the more consequential errors came in the form of inefficiency in conducting the examination and then communicating the wrong information —An act that could have dire consequences for the patient's health—. Because of this, it is appropriate to say that the examination was not at all efficient.

4.1.2.2. Feedback from RHP

The paramedic who was playing the role of the RHP gave the feedback below to the Control Simulation Number 1

-**How confident are you that you got all the information you needed to get a diagnosis** = 8/20 pts.

-**How confident are you that the examinations from the paramedic were done properly?** = 6/20 pts.

-**How easy was it to conduct the examination through the phone?** = 8/20 pts.

- **Any other comments about your experience?** "*It was easy to show them what to do because of good communication. But i [sic] had to trust them doing the right examination*".

The paramedic is scoring the items on the lower half of the scale. This describes a lower level of content with the questions expressed above.

He does comment on concern regarding having to trust that the PM is doing the examinations correctly. As seen in the previous sections, this concern is not unfounded.

4.2. Control Simulation 2

4.2.1. Overview of Simulation 2

The simulation starts. The PM begins the basic examination as expected. Unlike the previous participant, this paramedic performs part of the FAST examination before contacting the RHP. This makes an impact on the duration of the call.

This participant chooses to take the call on speaker and places the phone on the floor. Because of this, the RHP's voice can be heard by all the participants in the room. This allows the RHP to speak directly with the patient without the PM having to work as an interloper.
Furthermore, the PM is able to work hands-free, incorporating both hands into the examination.

During the examination, the RHP prompted the PM to carry out examination #2 in NIHSS (best gaze). When communicating the result of this to the RHP, the PM said the sign happened 'in the right eye'. In reality, the sign occurred on the left eye of the patient. It became obvious that because the PM was kneeling in front of the patient, they were describing right and left from their personal perspective, instead of the patient's. This is an important find because right and left sides may attend to their own specific symptomatology, especially in the case of a stroke. Confusing these sides may lead to misdiagnosis and or patient harm.

On step #7 in NIHSS the patient is supposed to extend their arms to the side and then — keeping the arms extended — touch the tip of their nose. During this roleplay the PM did not communicate these instructions to the patient, thus they did not extend their arms, and touched their nose while slouching, making the whole exercise moot. They were not instructed to correct her posture by the PM and the RHP assumed that the patient was doing the procedure correctly.

In the same way it happened in the previous simulation, the PM neglected to make sensory examinations (NIHSS step #8) on the lower parts of the body.

In summary, because there was no visual component, the RHP could not see that the PM made a mistake reporting signs on the wrong part of the body, and thus he was working with erroneous information to make his/her diagnosis.

4.2.2. Results

4.2.2.1. Efficiency of Examination

Could NIHSS be practiced correctly?

No. Step #7 (Limb Ataxia) and Step# 8 (Sensory) were done incorrectly. In Step #2 (best gaze) the exam was conducted correctly, but the information communicated was wrong (with the PM communicating problems with the right eye, when it was the left one which was presenting signs). In all these cases the RHP was communicated the wrong information.

Were there harmful Interventions? If so, how harmful were they

According to the reviewer (Paramedic Trainer), with three steps of NIHSS conducted erroneously, the RHP could have easily made the wrong diagnosis and indicate other actions to the patient other than the recommended path for a person with a posterior stroke case. This could result in severe damage to the patient.

Was the overall examination efficient?

While there were comparative improvements made in the efficiency of the examination due to the speaker/hands free communication, the fact that the RHP was unable to see and amend the parts of the examination which were incorrect, resulted in inaccurate information transmitted and consequently, the wrong decision of treatment could have been made for the patient. Because of this it is not possible to say that the examination process was efficient.

4.2.2.2. Feedback from RHP

The paramedic who was playing the role of the RHP gave the feedback below to simulation number 2.

-How confident are you that you got all the information you needed to get a diagnosis = 8/20 pts.

-How confident are you that the examination from the paramedic was done properly? = 6/20

-**How easy was it to conduct the examination through the phone?** = 7/20.

- A**ny other comments about your experience?** "*It was also here good communication. "Clear and nice voice."*

The RHP graded all questions under the lower half of the scale. The RHP was not able to notice that the examination was conducted wrongly. Their self-reported level of confidence was 6/20. This feedback reflects a low level of trust in the process and experience.

4.3. Control Simulation 3

4.3.1. Overview of Simulation 1

The PM began the examination. When the phone call portion of the process started the participant chose to put the phone on speaker. However, instead of placing it on a surface, as the participant in the previous simulation had done, this PM chose to hold the phone with one hand throughout the rest of the process. This limited his/her range of motion and affected the way he/she performed the examinations.

During the examination, the RHP prompted the PM to carry out Step #2 in NIHSS (best gaze). The paramedic conducted the examination quite fast and missed the sign that the patient was showing on their left eye. He/she communicated to the RHP that the patient's reaction was normal when this was not the case.

During Step #3 in NIHSS (visual field testing), the paramedic conducted the examination one-handed, when for this exam, it is necessary that both hands are used so that the patient receives the correct stimuli. Furthermore, the paramedic communicated a sign showing on the right side, when it actually belonged to the left side of the patient. In the same way it happened in the previous case, they were reporting on 'their side' (right) instead of the patient's (left).

Step #8 (sensory test) was not conducted in the lower part of the body.

The examinations were performed very fast. After the simulation ended, when interviewing the PM trainer who was evaluating the simulations, and the person playing the patient (who is a healthcare professional as well), both mentioned that all the examinations were done too fast, to the point where it was easy to see how the PM missed the signs they did.

4.3.2. Efficiency of Examination.

Could NIHSS be practiced correctly?

No. Step #2 (best gaze) was conducted too fast and the PM missed the sign the patient was portraying. In Step #3 (visual field testing) the exam was conducted one-handed. Furthermore, the PM confused the patient's left side with their right side and reported a sign on the wrong side of the patient's body. During Step #8 (Sensory) the PM did not observe the signs in the lower part of the body. In all these cases, the RHP was given the wrong information.

Were there harmful Interventions? If so, how harmful were they?

According to the reviewer, similarly to the previous simulation, on various occasions, the wrong information was communicated to the RHP. This would have resulted in the failure to recognize the posterior stroke the patient was presenting, leading to alternative treatments to be taken, instead of the time-sensitive measures needed in cases with such diagnosis. These could lead to severe patient harm.

Was the overall examination efficient?

This simulation was consistent with the issues of the previous two. There was poor execution of the exam, missed signs, and falsely reported information. Because of this, it is not possible to say that the examination was efficient.

4.3.3. Feedback from RHP.

The paramedic who was playing the role of the RHP gave the feedback below to simulation number 3.

-How confident are you that you got all the information you needed to get a diagnosis = 9/20 pts.

-**How confident are you that the examination from the paramedic were done properly?** = 8/20

-**How easy was it to conduct the examination through the phone?** = 7/20.

- Any other comments about your experience? *"I had the same experience here. I cant [sic] see the pasient [sic] or the paramedics so i[sic] have to trust them doing the right examination."*

The participant scored all three questions in the lower half of the scale.

Similarly to the previous cases, the PM was unable to discern if the PM did the examination correctly, reporting a low confidence level (6/20) over the accuracy of the examination.

----

5. Intervention Simulations (PreViS Group).

5.1. PreViS Group Simulation 1

5.2.1. Overview of Simulation

The simulation started similar to those in the control group.
When the phone call began, after an initial exchange, the PM requested the RHP if they could connect through the PreViS system. Upon the confirmation, the PM started to prepare the headset, which was already outside of the protective case, and turned on.
In the meantime, the PM explained to the patient what the headset was for, and asked them if they agreed to the device to be used for the examination. After the patient acquiesced, the PM put the headset on, made small adjustments, and started to make the call. As mentioned above, this process lasted just under a minute.
The connection was made in a straightforward manner.

The examination started. The headset was on speaker mode, thus the patient could hear the RHP themselves: When the RHP asked a question to the patient, they heard it, and responded directly to the RHP through the headset.
e.g.
RHP: *"How old are you?"*
Patient: *"53"*.
RHP: *[proceeds to the next question]*

Furthermore, the RHP would dictate directly to the patient how to do certain parts of the examination without needing the interloping from the PM (e.g. "*extend your arms to the side... Now close your eyes and touch the tip of your nose with your right hand*").

Overall The NIHSS examination went through in a fluid manner.

During step #2 (Best Gaze), the RHP was able to see the sign that the patient was portraying directly on the camera.

In Step #3 the RHP corrected the PM when they tried to conduct the examination one-handed. The RHP then instructed in more detail how to conduct it, until the PM did it correctly. After that, the RHP was able to 'catch' the sign that the patient was portraying.

During the examinations where the area of the patient the RHP needed to see was 'off camera' or away from the frame, the RHP instructed the PM to '*adjust the camera a little bit*', or '*go further down*' etc. These interactions happened twice during the conversation and were resolved in less than 4 seconds each.
The RHP skipped one of the steps in the examination by mistake, however, he remembered it later during the examination and completed it.

The rest of the examination went through without any relevant incidents.

5.2.2. Results PreViS Group Simulation 1

5.2.2.1. Headset Interactions: Technical Review.

Were there any instances of user error?

It can be argued that forgetting to notice that the relevant part of the patient was not on camera can be classified as an user error, which is pointing towards a mismatch between the user's mental model and the system's output.

Were there any delays or problems starting connection?

No.

Were there any instances of voice-command failures?

All the commands spoken by the PM were taken by the system correctly on the first try. However, in one instance the head-camera picked up on a random word from the RHP and took it as a command. This is a significant find, as this vulnerability may lead to an incident. This scenario was repeated more severely in the simulations described on the following pages.

Were there disturbances in Audio (Lag, interruptions, problems with volume, etc).

No instances recorded.

Were there notable fidgeting and readjustment instances with the headset?
During the 9+ minutes of examination, there were 6 instances where the PM placed his hands on the device to make minor adjustments. All but one of the instances lasted less than 2 seconds each. Body language did not indicate large gestures of discomfort.

5.2.2.2. **Efficiency of Examination**
Could NIHSS be practiced correctly?
Yes. All the stages of the examination were done correctly, confirmed by the standards of the RHP and PM reviewer.
Were there harmful Interventions? If so, how harmful were they
No harmful Interventions detected.
Was the overall examination efficient?
Given the following: ● All the signs from the patient's condition were easily seen on camera by the RHP. ● NIHSS was performed successfully, and there were no harmful interventions. ● Any missteps from the PM were recognized and easily corrected. ● The examination was done in a fluid manner, and there were no signs of the PM struggling with the headset. ● There were no technical difficulties recorded. ● There was positive feedback from the RHP (see next section). **It is possible to say that the examination was efficient**.

5.2.2.3. Feedback from RHP

The paramedic playing the role of the RHP gave the feedback below to the PreVis simulation number 1.

- **Did you experience any miscommunications?**
  "No. It was easy to communicate here because of good video and audio. I lost one question but was my fault and i [sic] took it in the end".

- **Was there ever a drop in quality of video or audio that got in the way of examinations.**
  [No answer in this field. Upon reviewing the footage transmitted from the headset, there was no drop of quality found].

- **Was the image of the patient in focus enough so that you could see what you needed from the patient?**
  "Yes. I could see the patient all the time. I had to tell them a couple of times to move the camera when doing examination"

- **How confident are you that you got all the information you needed to get a diagnosis** = 17/20 pts.

- **How confident are you that the examination from the paramedic was done properly?** = 18/20

- **How easy was it to conduct the examination through the headset? =** 16/20.

- **Any other comments about your experience?** "*This was a good experience"*

When asked follow-up about the reason behind the 16/20, the RHP mentioned the need to ask the PM to switch the view point so they could see what was needed (i.e. sometimes the part of the body of the person was out of frame). This topic will be addressed again in the following simulations.

5.2. PreViS Group Simulation 2

5.2.1. Overview of Simulation

During this simulation the PM verbally expressed that he/she was very nervous. This led them to hesitate throughout several steps of the process.
In the early stages of the roleplay (before the phone/headset part of the examination he/she broke the fourth wall of the roleplay looking for assurance/approval of their actions). The PM was given no feedback from the observers of the roleplay. He/she discussed their plan of action with the Partner-PM. Together they agreed to call the RHP.

After that, he/she started the process of making the call with the headset. The PM broke the 4th wall once more to ask about how to turn on the modem. He/she was given quick instructions by the partner Partner-PM and started it in a matter of seconds.

The PM briefly asked the patient if they could film her with the head-camera.

The initial commands were accepted in the first try ('start app'; 'ring test-room').
However, within the first minute of the conversation, the zoom command was activated even though no such command was uttered. The RHP had to ask the PM to adjust the image. The PM had to readjust their position and utter the command 'Zoom' to make sure the patient image was still well framed on the optical microdisplay feed.

After this, two more zoom commands were activated without user input. On three occasions the 'light' (flashlight) command activated unprompted. Observing their body language, one could see that this was very distracting for the PM.

The PM made small adjustments to the microdisplay. It was visible that he/she was concentrating on looking directly at the microdisplay --closing one eye--, and making sure that the right areas were in focus.
The zoom command was activated by itself once more. The time after that, the command zoom was prompted by the paramedic successfully.

At one point, the patient got instructed by the RHP to step #5 of NIHSS examinations (motor function arm). The RHP could not see the full figure of the patient, so they asked the PM to stand back. The PM then switched from kneeling to a standing position in order to fit the full image of the patient into the frame he/she was transmitting to the RHP. The PM could once again be seen closing one eye to focus on the image he/she was transmitting on the optical microdisplay.

Another example on the interaction of the RHP asking the PM to adjust the camera to frame the area they needed to see happened during NIHSS step #4 (facial palsy), where it was necessary for the camera to capture the patient's face.

RHP: [to the patient] *can you smile please*?
[Patient smiles]
RHP: [to PM] *can you move the camera a little bit?*
[PM adjusts the camera].
RHP: [seeing the required image on the frame] *Good*.

During one of the examinations the RHP prompted the PM to zoom or focus the image on a part of the patient, and he/she does, by using the zoom command. This command was accepted on the first try.
Further along the process the RHP asks the PM to 'zoom in' again. In this case, however, the 'zoom' command was already in 'zoom in' mode, and both the PM and RHP realized this when the command was accepted and reverted to zoom out. This command was not taken the first, but the second time it was uttered.

The review of the internal video that was transmitted from the headset revealed a design flaw in the equipment: the adjustable arm the optical microdisplay is on —which can be moved to fit the users needs— can inadvertently be placed in ways that can cover part of the camera. As it can be seen on the images below (lower left corner), this can result in occlusions of the image transmitted to the RHP.
Despite the intermittent technical interruptions, the RHP was able to get all the relevant signs to do a proper examination on camera. Much like the previous simulation, the RHP took the lead guiding the process. On one of the examinations he saw how the patient was not holding the right posture. They then corrected the patient themselves until the correct position was performed.

The simulation finished without further notable incidents.

A Post-simulation interview with the participant who played the patient revealed that their experience with this paramedic was that they seemed to be so preoccupied with dealing with the headset, that they (the patient) felt somewhat neglected at times. They mentioned that it was easy to imagine how in this scenario, a patient could feel alienated from the process.

5.2.2. Results PreViS Group Simulation 2

5.2.2.1. Headset Interactions: Technical Review.

Were there any instances of user error?

Yes, two instances.
The same issue where the PM did not notice that the transmitted image was not aligned with their POV happened in this simulation. The PM had to be reminded that they had to make sure they were framing the relevant parts of the patient on camera.

While adjusting the microdisplay, the PM inadvertently placed the arm of the microdisplay in a way where it covered a small portion of the camera. Because of this there was a small amount of occlusion on the feed the RHP was receiving. Much like the framing issue described above, these can be arguably not defined as user error and more a design flaw of the equipment (see Discussion section).

Were there any delays or problems starting connection?

No.

Were there any instances of voice-command failures?

There was one instance where the user voiced a command and the system did not recognize it on the first try.
On the other hand, out of 14 commands recognized by the system only 8 of them were real/intentional commands spoken by the user.
It is unclear what triggers these random activations of the zoom and light command. It's important to note that even though the examination was successful, these occurrences were definitely disruptive to the process. The zoom function getting triggered randomly definitely affected what the RHP could see on their screen. It also burdened the Paramedic with the load of having to stop the process and having to reposition themselves because the system was not cooperating. Furthermore, it is easy to see how having the light turned on at random times can generate frustration and or negative feelings from both the PM and the patient.

Were there disturbances in Audio (Lag, interruptions, problems with volume, etc)?

No instances recorded in the audiovisual transmission.

Were there notable fidgeting and readjustment instances with the headset?
The adjustments that were done in this case had less to do with the 'fit' of the headset, and more with the adjustment of the camera itself, and how to make the desired picture get into frame. There were 7 instances of adjusting the microdisplay. 4 were resolved in under 3 seconds, but three of them were in the 9, 12 and 30 seconds range. There was a pronounced effort on the paramedic's part to make sure that he/she was 'broadcasting' the correct thing. He/she spent a total of 1 minute, 1 second doing adjustments[1] In some of these occurrences, the PM was seemingly somewhat strenuously keeping one eye shut in order to focus on the image on the optical microdisplay. Some of these adjustments were at the behest of the RHP, while some of them were unprompted. This indicates a ‘new task’ the PM had to add on their list of activities during the examination. From their feedback, the RHP also experienced certain levels of frustration at the technical difficulties. For professionals that work on very tight schedules —as it will be the case for many RHP— this may be frustrating, and may threaten adoption of the system.

5.2.2.2. Efficiency of Examination

Could NIHSS be practiced correctly?
Yes. All the stages of the examination were done correctly, confirmed by the standards of the RHP and PM reviewer.
Were there harmful Interventions? If so, how harmful were they
No harmful Interventions detected.
Was the overall examination efficient?
It is harder to come to a conclusion in this simulation. Was the diagnosis properly achieved? Yes. According to the feedback of both the RHP and the reviewer, they saw the information they needed to see to make a proper diagnosis. This can be confirmed by the video captured by the headset, as all relevant signs were captured by the camera. However, the simulation was marred by the glitches in the headset and the effects of this had their consequences

on the PM's performance and effort, the RHP's ease of access to the required information, and the patient's feeling of detachment from the PM's care. While these may not easily fall under the life threatening consequences in this particular simulation, it definitely negatively affects the user's experience on the main members of this ecology. Because of this, one cannot say the process in this simulation was efficient.

5.2.2.3. Feedback from RHP

The paramedic who was playing the role of the RHP gave the feedback below to the PreVis simulation Number 2.

Did you experience any miscommunications:
[answer left blank]

Was there ever a drop in quality of video or audio that got in the way of examinations.
"The quality of the video was good".

Was the image of the patient in focus enough so that you could see what you needed from the patient?
"The focus on the patient was not so good all the time[2]"

Any other comments about your experience?
"*It was sone [sic] technical problems with the zoom on the camera.*
and the paramedic couldn't find the right position for the focus when doing examination".

Scales Section
The participant scored all three questions in the upper quarter of the scale.

- **How confident are you that you got all the information you needed to get a diagnosis** = 16/20 pts.

- **How confident are you that the examination from the paramedic was done properly**? = 16/20

- **How easy was it to conduct the examination through the headset?** = 15/20.

- **Any other comments about your experience?**:

[2] It is important to reiterate that when speaking of the patient being 'in focus', this is referencing having the desired image of the patient sufficiently 'in frame', rather than referring to blurriness on the image. The internal video captured by the headset was examined to confirm that the latter was not the case.

"It seem to bee [sic] some problems from the user here. It took a little to [sic] much time doing the camera ready"

5.3. PreViS Group Simulation 3.

5.3.1. Overview of Simulation

This examination went a lot smoother than the previous one. In both terms of PM and device performance.

The PM asked for consent for the patient to use the headset and then proceeded to prepare the device by pulling it out of the pouch. The preparation time from first contact with the device to initiating the examination was shorter than the previous examination.
The examination took a pretty fluid course. The 'phantom command' issue experienced on the previous simulation --where the system implemented a command without user input-- only happened once, early during the call, when the zoom was seemingly activated by itself. The PM noticed it as it happened, and reverted it to the previous state. The occlusion in the corner of the image from the previous simulation was also present this time. There were no other device related interruptions for the duration of the simulation.

During NIHSS step #3, the RHP realized that the PM was not conducting the examination accurately and corrected him/her.

During one close-up examination the RHP said "*can you move back? I can't see the patient*". The PM then changed positions and checked the microdisplay to see if the patient was properly framed in the shot.

From then on, the PM was more aware of framing the patient. During NIHSS step number #5, he/she moved to encompass the person unprompted, so that the PM could see that there was no drift on the arms. He/she did the same in the case of the leg-based examination.
The rest of the examination went on without any incident to report.

5.3.2. Results PreViS Group Simulation 3

5.3.2.1. Headset Interactions: Technical Review.

Were there any instances of user error?

Only one instance of having to be reminded by the RHP to readjust their position / framing. After this, the participant was very aware of framing the patient correctly

Were there any delays or problems starting connection?

No.

Were there any instances of voice-command failures?

There was only one 'Phantom Command' in the beginning.

Were there disturbances in Audio (Lag, interruptions, problems with volume, etc).

No instances recorded.

Were there notable fidgeting and readjustment instances with the headset?

Despite the readjusting made to fit the patient in the frame, the PM never touched the arm holding the microdisplay or the headset itself after the initial setup stage. All the 'framing' was done with posture changes (i.e. slight movements of the neck, and/or changing from a kneeling position to a standing one).

5.3.2.2. Efficiency of Examination

Could NIHSS be practiced correctly?

Yes. All the stages of the examination were done correctly. This was confirmed by the standards of the RHP and

PM trainer.
Were there harmful Interventions? If so, how harmful were they?
No harmful Interventions detected.
Overall, was the process efficient?
Given the following: ● All the signs from the patient's conditions were easily seen on camera by the RHP. ● NIHSS was performed successfully, and there were no harmful interventions. ● Any missteps from the PM were recognized and easily corrected. ● The examination was done in a fluid manner, and there were no signs of the PM struggling with the headset. ● There was only one 'phantom command' in this simulation, and it was remedied within seconds. ● There was positive feedback from the RHP (see next section). It is possible to say that the examination was efficient.

5.3.2.3. Feedback from RHP

The paramedic who was playing the role of the RHP gave the feedback below to the PreVis simulation number 3.

Did you experience any miscommunications?
"*Very good communication this time*"

Was there ever a drop in quality of video or audio that got in the way of examinations.
"No. Very good quality".

Was the image of the patient in focus enough so that you could see what you needed from the patient?

"*Yes. This time [sic] i could see all the examination*"

Any other comments about your experience?
"*This show mee [sic] that this is a good method to examinate [sic] when the paramedic is doing the right things with the zoom. It took a little to [sic] much time to get the camera ready*".

Note: When the RHP says that "*the PM took a little too much time to get the camera ready*" it is relevant to mention that this PM was slower at the preparation time when compared to the first simulation —where preparation time was shorter due partly to the fact that the PM showed up with the headset out of the case and turned on—, but faster in comparison to the previous simulation.

Scales:
The participant scored all three questions in the upper quarter of the scale.

● **How confident are you that you got all the information you needed to get a diagnosis** = 18/20 pts.

● **How confident are you that the examination from the paramedic was done properly?** = 19/20

● **How easy was it to conduct the examination through the headset?** = 19/20.

● **Any other comment about your experience?** *"Its best to examinate [sic] the patient directly through the camera."*

---

6. Results from GUI Usability Inspection

GUI-Specific Issues

The following is a description of the results observed during the usability inspection PreVIS Graphic Interface and navigation.

Issue: GUI Navigational Flaws: Lack of Navigational Cues

When observing the PreVis digital interface, it can be presumed that one of the biggest issues challenging the design of this GUI is the very limited quantity of real estate available. However, as it will be seen on some of the following images, more than suffering from a lack of space, the interface is struggling with the misuse of it.

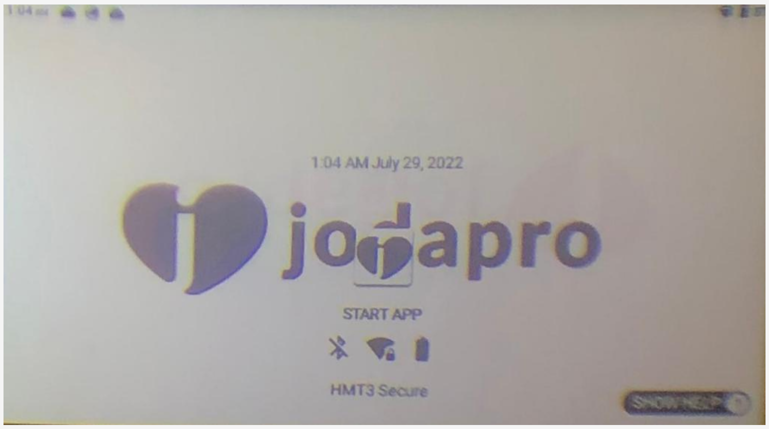


Fig. A.01. Home Page UI
This is the first screen the user can interact with.

An example of this is on the 'Homepage' of the app. This is the first screen the user can interact with when turning on the device. As it can be seen, there are no obvious navigational elements that can help the user create a mental model of the information architecture of the place, and/or how to navigate through it. The element with the highest visibility is a branding background of JodaPro, followed by a central (non functional) 'button' with the same logo. Under it there is a simple text that says 'Start App'. To the right hand corner there is a button with a different graphic language which says 'Show Help'.

Saying the command 'Start App' is the first step to get to the primary functional section of the system, which is the interface where the PM can call a RHP. This section has been given the name 'Kompetanseavdeligen'.

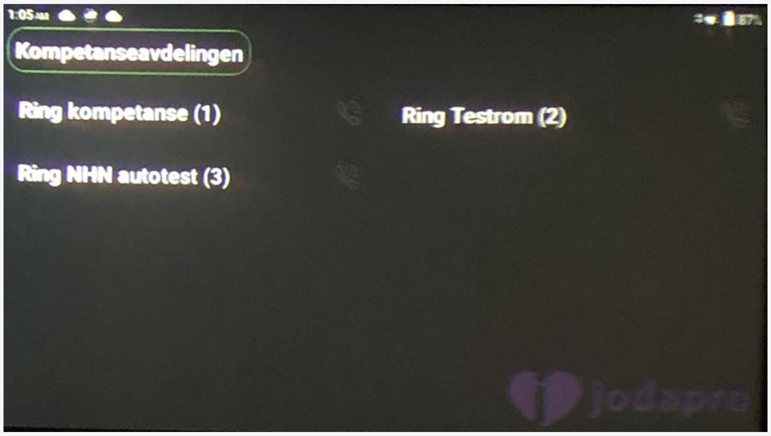


Fig. A.02. 'Kompetanseavdelingen' UI.

This takes the user to the Kompetanseavdelingen section of the system. This opens the call log screen where users can select the number they want to video-call to. Voice Command: Start App

'Kompetanseavdeligen' suffers from the same lack of navigational elements as the Home Page. As it can be seen on the image, this is not due to the lack of space either. One way to highlight how this configuration is hurting the ease of navigation for the user is to perform an abbreviated version of an Instone test.

The reason why it has been abbreviated is that some questions from the Instone Test are arguably not relevant to PreViS, as this questionnaire is dedicated towards less specific projects than this, i.e. consumer digital products.

In the section below eight questions from the Instone Test are asked from the Kompetanseavdeligen' interface.

6.1. Instone Test

Instone Test Prompt	Answer
What is this page about?	The title 'Kompetanseavdeligen' does not do a very good job on describing the function of this page. Fortunately, the options below are written as 'Calls to Action' (CTAs) which if said out loud by the user will trigger the expected action (for example 'Ring Testroom' triggers a call to the Test Room environment).

	Recommendation: Use a title more relevant to the section, and perhaps a short descriptive meta text to let the user know that they can read the commands to trigger the call.
What are the major sections of this site?	There are no navigation elements in this interface which can give context about the other sections of the site. **Recommendation:** Implement some type of graphic cue (heading, et al) which denotes the name of this section. Whenever possible align the voice commands to match the names of said sections (for example, saying 'Home' should trigger navigation back to Home). In this way the graphic presence of the names of sections in the navigational menus can also work as reminders of the commands the user must say to trigger the navigation to these destinations (this is an implementation of the 'Recognition over Recall' heuristic).
What major section is this page in? **What is "up" 1 level from here?** **How do I get to the top of this section of the site?**	It is unknown by the GUI alone where in the hierarchy is this section and/or what is 'up from here'. There are no breadcrumbs or other elements which can clue the users about this. Without knowing the necessary command by memory, one would have to guess how to move to other levels in the hierarchy. **Recommendation:** Use breadcrumbs or another graphic cue to communicate this. See second recommendation on the last question.
How might you get to this page from the site home page?	The user would have to read the phrase 'Start App' which is shown on the homepage on a label to arrive at the 'Kompetanseavdeligen' section. While the non-sequitur nature of the taxonomy will not necessarily make the user get lost, this is definitely improper use of navigation, and a violation of the Consistency Heuristic: e.g. calling the name of a command and ending on a section of a different name. **Recommendation:** As previously addressed, one should associate the commands that will take the user to a section to the name of the section itself: for example the graphic cue should say [Name of Section] and the

	command should be Navigate to [Name of Section], and/or [Name of Section] itself.
What site is this?	The name of the system (PreViS) is nowhere to be found on the interface. This may be a lesser sin in the context of PreViS being a working tool, and not a consumer product. The impact of this is lower in comparison with other GUI issues on this list.

As it can be seen in the previous page, even with the system's very shallow hierarchical depth, the current GUI lacks the navigational cues to make the user navigate easily.

In an average personal computer interface, where the user can ‘click around’, and through trial and error understand which links go where perhaps the cognitive load of exploring such a shallow hierarchy might be less taxing than staring at the screen and trying to conjure up the correct commands to move to another destination.

The Existing Menu.

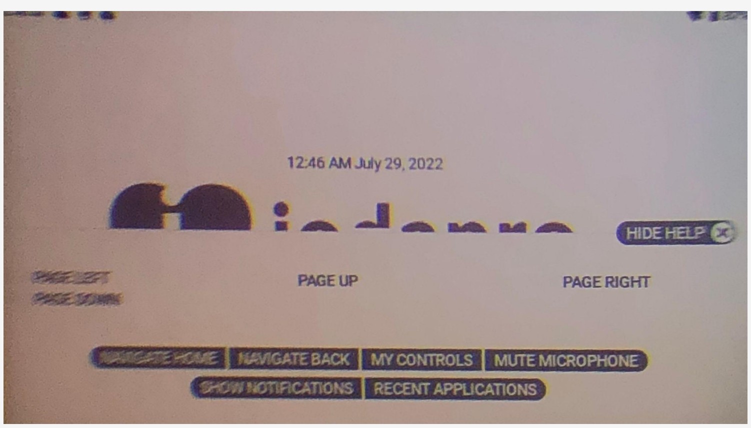


Fig. A.03. ‘Show Help’ UI

This shows a menu with different options (Voice Command: Show Help).

Returning to the topic of the Homepage, calling out the 'Show Help' command unveils a menu which mixes some navigational commands (Navigate Home, Navigate back, My Controls, Recent Applications) with functions (mute microphone, recent applications).

This menu is interesting for several reasons: First, it is not related to the PreViS direct functions: there is no button for accessing Kompetanseavdeligen, which is the entry point into the main functionality of PreViS i.e. Calling an RHP; 'Show Notifications' and 'Recent Applications' have currently no relevant use for PreViS activities. Second, 'Navigate Home' and 'Navigate Back' do the same action in this context, which is simply closing the menu and 'bearing' the homepage again. Third, this menu is not available on the 'Kompetanseavdeligen' where it may arguably be more useful to navigate back home with. It is also graphically missing in the 'Call Space', although if the user remembers to utter the command ‘Show Help’, the menu will appear. In this case, they would have to recall that this option is available, despite the lack of graphic cues to indicate it.. This is yet another violation of the 'recognition vs. recall' heuristic.

This menu is proof that navigational elements can appear satisfactorily on the screen. However it needs to be reconceived.

Recommendation: Rewrite this menu to have the three main levels of the hierarchy: Home, Kompetanseavdeligen, and My controls. If necessary, it can also hold a few voice-command shortcuts to needed functions (more user-data is needed to understand what these should be, if any). As previously mentioned, the menu should always be visible throughout the system so that recognition over recall is always fulfilled

Call Room space.

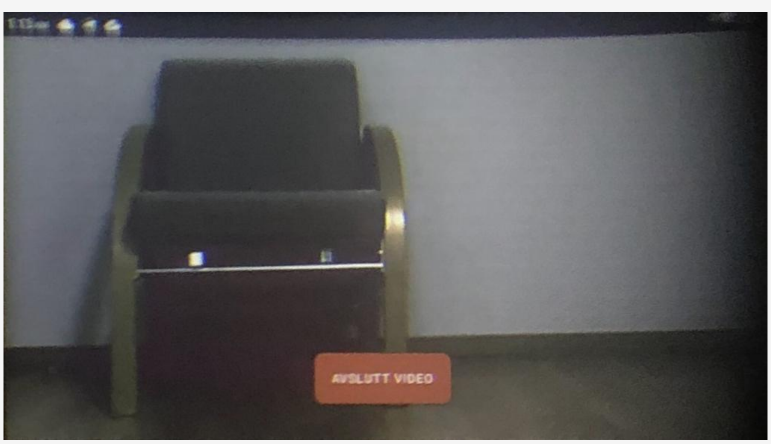


Fig. A04’ Call Space’ UI.

Users arrive at the ‘Call Space’ where video calls happen (Voice Command: Ring [Selection). In this section the microdisplay shows what the camera is broadcasting, and a button which says ‘Avslutt Video’, which allows the user to terminate the call. In this space users can use the commands: Zoom, and Light to trigger enhancement on the camera, and the ignition of the integrated LED flashlight, respectively.

The 'Avslutt Video' (in English 'Exit Video') button shows an acceptable implementation. It is an explicit call to action which when said by the user triggers the expected outcome (i.e. hanging up the call).

Some important graphic elements missing from this screen are related to the 'Zoom' and 'Light' function. The reasons for this are Voice-Command related (to fulfill the recognition instead of recall heuristic) but perhaps equally importantly, to indicate whether the camera's zoom is currently on or off. This comes from the observations in the simulations, where users called on the zoom command without noticing it was already on. When this happened the zoom was 'turned off' and the camera reverted back further away from the original 'shot', giving the user the opposite outcome of their requirement.

Recommendation: To fix this error it is necessary to use graphic cues on the UI to show when the 'Zoom' and 'Light' are On or Off. The implementation of this would fulfill the heuristics of Visibility of System Status, as well as Cognition Facilitation (Wheeler Atkinson et al., 2007).

Depurate Irrelevant Navigational Sections.

If we consider that the true entry point to start using PreVis is the Kompetanseavdeligen, then there is merit to the question: What is the point of the HomePage in such a shallow hierarchy? It would be interesting to explore iterations where the navigation structure of the HomePage includes the main functionalities of Kompetanseavdeligen; making this one of the first thing the user sees when they turn on the device, bringing the user one step closer to accessing their primary goal (calling the RHP). Testing and iterating is necessary to understand if this measure is more useful to the user than the alternatives.

6.2. Interpreting the SUS Score.

What constitutes an acceptable SUS score?

It is important to understand what are the metrics that would point to an acceptable SUS score.

According to (Bangor, A et.al, and their experience of analyzing a decade's worth of SUS data, a so-called 'university grade analogue' is the most (accepted way) of grading the products.

> *"[The university grade analogue] is well supported by both the adjective rating scale data and our own collective experience that takes into account the nature of the failures. This means that products which are at least passable have SUS scores above 70, with better products scoring in the high 70s to upper 80s. Truly superior products score better than 90. Products with scores of less than 70 should be considered candidates for increased scrutiny and continued improvement and should be judged to be marginal at best".*
> (Bangor, A et.al 2008).

The score in the PreViS evaluation was 80. According to Bangor's et.al. assessments, this means the usability of the system is perceived as upwards of passable, in the 'better products' echelon. These results are however a simple point of reference, and further research is needed for validation due to the following reasons:

1. The low sampling number makes the dataset quite weak.
2. The fact that the researcher had no direct control over the recruitment of participants, and they are foreing to the paramedic community means that there are possible bias blindspots, i.e., there is the possibility that these responses might have been affected by social and/or power dynamics within the organization. Future research would benefit from understanding and controlling for conditions such as these.

In order to understand the impact of these scores, it is relevant to see them transposed through other methods of observation. As Bangor's et.al. mentions:

> "*As with any metric, the SUS score should not be used in isolation to make absolute judgments about the "goodness" of a given product. Factors such as success rate and the nature of the failures observed when the system was tested with representative users should play a large part in determining how usable a product is (ISO, 1998)*"
> (Bangor, A et.al 2008).

7. Training-Specific Recommendations:

The following is a list of suggestions to integrate in PreViS related training sessions. These come as a result of the observations in different training sessions as well as the simulation central to this paper. The intention of these is to complement the learning curve of the user.

Preparing the Headset during the ride:
It was seen in one of the simulations that if the participants prepare the headset in the ambulance while on route to the emergency, they can save the preparation time at the scene. Incorporating this learning into the training routines will create a more streamlined interaction with the patient and a faster connection to decision support (RHP).

Discovering Eye-Dominance.
The manual for the headset states that users must understand which is their dominant eye before deciding on which side of their face to place the microdisplay. Said manual proposes a simple exercise to self-diagnose on which is one's dominant eye. This exercise was not implemented in the training sessions the researcher attended. Users seemingly chose the side to put the microdisplay on depending on whether they were right or left-handed. Eye-dominance and hand-dominance are not always the same. Using the microdisplay on the correct side might make the visual interaction with the microdisplay a more accurate and pleasant experience for the user. Because of this, it is recommended to integrate the eye-dominance exercise on early sessions of the training.

Standardize the Request for Consent.
There is currently no official way to request consent from the patient on whether they agree to be filmed, broadcasted to a third party they cannot see. All three paramedics in the simulation used different phrasing, with different levels of details and disclosures. Interviews with paramedics who have used the system with real life patients expressed that they have never experienced negative feedback from them, (if anything, the novelty of the device has been received very positively). That being said, it is worth designing the more apt, standard prompt to elicit consent which can cover all the bases and puts the patients at ease. Further research is needed to understand esåecifically what this interaction should encompass. The overall goal is that no lack of knowledge, or source of patient anxiety is falling through the cracks.

Train to Avoid Patient-Alienation.

After the simulations, interviews with the participant who played the patient revealed that one of the participants seemed to be so preoccupied with dealing with the headset, that it was easy to imagine how a patient could feel alienated from the process. Understanding that this is a possibility is the first step to address it. Training sessions could include this as a factor to be aware of, allowing the paramedics to be able to understand it, and practice techniques on how to avoid it.

Recognize Risky Patients

In one Brain-Writing session conducted with paramedics, participants were asked: When do you think using PreViS is a risk? Paramedics named a list of patient-types where they considered that the presence of PreVis could make situations more acute. All mentions fell under the categories of Altered-State cases, being either because of psychiatric reasons and/or substance abuse. It would be important to address these risks during training sessions, including guidance on how to assess and when not to use PreViS.

The Importance of Project Ownership and User Feedback.

This project has been developed from paramedic by paramedics. This is an important factor that should persist. Use training sessions to instill a sense of ownership on the paramedics who are being trained to use the system. In the hyper-specific context PreViS is used in, their active observations, engagement and feedback are one of the most invaluable factors towards the development of the product. Make an effort to bypass power dynamics in the pursuit of honest and constant user feedback. Provide methods (preferably anonymous) where users can provide their experiences or suggestions on how to make the system better.

Collect feedback from RHPs on the quality of the Interaction

In the same vein as the topic above, involve RHPs in the process. Collect their feedback and experiences, as they are also an important part of the interaction, and hold a unique point of view on it. Consider there might be needs from that user group which are falling through the cracks.

8. NIHSS Form (detail)

NIH stroke scale

Admission date: Time:

Item	Score	Description	Side	Before	2 h	24 h	7D/ Disch
1a.Level of consciousness	0	Alert					
	1	Not alert, but arousable with minimal stimulation					
	2	Not alert, requires repeated stimulation to attend					
	3	Coma					
1b. LOC questions *Ask patient the month and their age*	0	Answers both correctly					
	1	Answers one correctly					
	2	Both incorrect					
1c. LOC commands *Ask patient to open/close eyes and form/release fist*	0	Obeys both correctly					
	1	Obeys one correctly					
	2	Both incorrect					
2.Best gaze *Only horisontal eye movement*	0	Normalt					
	1	Partial gaze palsy					
	2	Forced gaze palsy					
3. Visual field testing	0	No visual field loss					
	1	Partial hemianopia					
	2	Complete hemianopia					
	3	Bilateral hemianopia (blind, incl. Cortical blindness)					
4. Facial palsy *Ask patient to show teeth or raise eyebrows and close eyes tightly*	0	Normal symmetrical movemeent					
	1	Minor paralysis (flattened nasolabial fold, asymmetry on smiling)					
	2	Partial paralysis (total or near total paralysis of lower face)					
	3	Complete paralysis of one or both sides (absence of facial movement in the upper and lower face)					
5. Motor function arm	0	Normal (extends arm 90° or 45° for 10 sec without drift)	**Right**				
	1	Drift					
	2	Some effort against gravity					
	3	No effort against gravity	**Left**				
	4	No movement					
	9	Untestable (joint fused/limb amputated) *(do not add score)*					
6. Motor function leg	0	Normal (holds leg in 30° position for 5 sec without drift)	**Right**				
	1	Drift					
	2	some effort against gravity					
	3	No effort against gravity	**Left**				
	4	No movement					
	9	Untestable (joint fused/limb amputated) *(do not add score)*					
7. Limb ataxia	0	No ataxia					
	1	Present in one limb					
	2	Present in two limbs					
8. Sensory *Use pinprick to test arms, legs, trunk and face, compare side to side*	0	Normal					
	1	Mild to moderate decrease in sensation					
	2	Severe to total sensory loss					
9. Best language *Ask patient to describe picture, name items*	0	No aphasia					
	1	Mild to moderate aphasia					
	2	Severe aphasia					
	3	Mute					
10. Dysarthria *Ask patient to read several words*	0	Normal articulation					
	1	Mild to moderate slurring of words					
	2	Near unintelligible or unable to speak					
	9	Intubated or other physical barrier *(do not add score)*					
11. Extinction and inattention *Use visual double stimulation or sensory double stimulation*	0	Normal					
	1	Inattention or extinction to bilateral simultaneous stimulation in one of the sensory modalities					
	2	Hemi-inattention, severe or to more than one modality					
12. Distal motor function *Ask patient to extend his/her fingers as much as possible*	0	Normal	**Right**				
	1	At least some extension after 5 sec but not fully extended					
	2	No voluntary extention after 5 sec	**Left**				
		Total score:					

Fig. A.07. NIHSS Scale form which details the steps of the examination. From www.nihstrokescale.org, (n.d.)

NTNU

Norwegian University of
Science and Technology