



## User Study validates a Computer-Supported Integration and Video Management System for the OR

Conducted with 26 clinical users from 7 different hospitals

As the digitalization of healthcare progresses, imaging systems in operating rooms are becoming ever more numerous and complex. In light of the growing complexity, the respective software and hardware components must be user-friendly to facilitate patient-centric workflows in the OR. With this in mind, EIZO GmbH tested the intuitive operation of its Caliop software and Surgical Panel from the CuratOR product line. For decades, EIZO has been developing video monitor and display solutions for sectors ranging from industrial and automotive to air traffic control and medical applications.

In the healthcare sector, EIZO hopes to support the surgical staff by focusing on improving the efficiency and quality of services in particular. A key initiative is to integrate the growing number of imaging systems used in the OR onto a central hardware and software platform. This white paper highlights the results of a study carried out in 2018/2019 in cooperation with M3i GmbH. With 26 clinical experts from 7 hospitals, the study evaluated the new CuratOR Caliop integration, video management system and Surgical Panel hardware.

# EIZO

## CuratOR Caliop

EIZO designed its software solution for the operating room to support seamless interaction among all video components and the ability to integrate third-party components. CuratOR Caliop ensures the smooth routing of connected image systems within the OR, regardless of the manufacturer. The system is controlled quickly and easily through an intuitive, central user interface, which manages video routing and is precisely aligned with the OR workflow.

**With CuratOR Caliop, EIZO has developed software that supports multiple workflows on a central hardware platform, the CuratOR Surgical Panel.**

---

### Key Data

- 96% of the experts had a positive overall impression of the system
  - 100% of subjects asked found the system easy to learn
  - 100% of the experts asked felt confident working with the system
  - 40% reduction in task performance time after only one usage
  - The system achieved a SUS score of 81.1 out of 100 possible points, representing very good to excellent usability
-

# Scope of the User Tests

In conjunction with the M3i industry-in-clinic platform, multiple usability studies were performed using a range of test scenarios and test subjects from different specialties. The objective was to evaluate the user-friendliness and intuitiveness of the system's software and hardware components. This white paper presents the key conclusions from these studies.

James Berge, Product Manager at EIZO GmbH stressed that a quick and simple overview of all relevant information can be critical to the success of a surgical intervention. "This is why we are always interested in finding out what potential difficulties might occur when using our hardware and software in the OR environment. By better understanding such issues, we can optimize our products and minimize potential risks in their use" said Berge.

The objectives of the usability tests were to identify potential usage problems within the software, to verify marketing messages, and to evaluate and implement improvements based on feedback and observations of the participants' use of the CuratOR Caliop software in a simulated operating room environment. For the purposes of this

study, various user groups from the clinical field were questioned, including surgeons, OR nurses, and IT administrators. The results of the latter study will be analyzed separately and are not included in this white paper. The usability test examined the following aspects:

- Simple and intuitive menu navigation
- Seamless integration of signals from multiple imaging systems
- Local operation using a mouse / keyboard, touch screen and remote control, for example using a footswitch
- Routing multiple image sources using a single user interface
- Confirmation of the functional scope of the extended CuratOR Caliop vm and Caliop doc modules







# CuratOR Caliop – Basic Functions

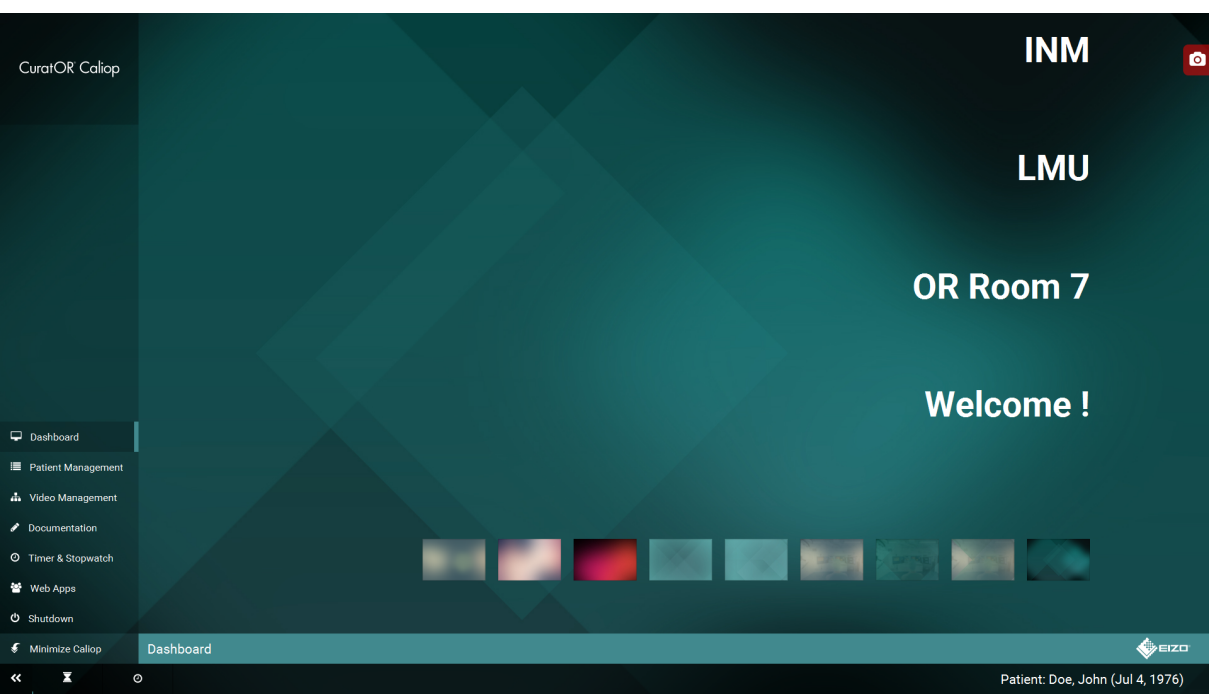
**Caliop comes equipped with a range of basic functions to support workflows in the OR.**

## Dashboard

The dashboard represents the starting point within an OR. It enables the selection of preferences such as the design of the user interface and offers an overview of current OR information such as active video streams and recordings or running timers. The dashboard is intended as a central location where configuration options for individual users and user groups can be set.

## Transparency

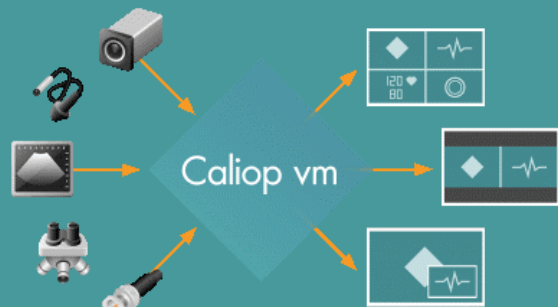
The functionality of the various Caliop modules is accessed through the main menu, where tasks are sequenced in accordance with a standard OR workflow. Active functions such as recording, streaming, and running stopwatches and timers are displayed clearly and can be registered at a glance, regardless of which module the user happens to be working in.





## Video Management (Caliop vm)

- Control of video from multiple connected devices
- Routing of video content to displays, recording and streaming devices
- Layout arrangement of the multiple images
- Visual display of routed connections and running functions such as active video recording and streaming

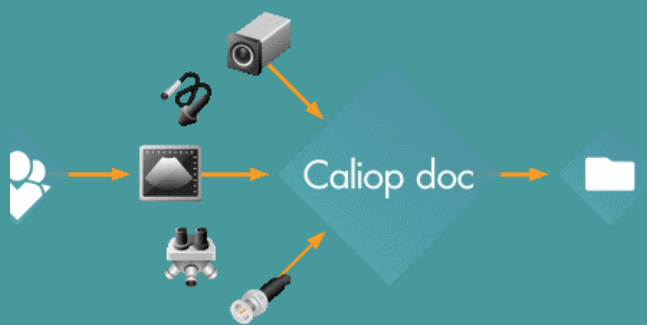


## Documentation and Archiving (Caliop doc)

- Starting and stopping recording of video sources
- Creation of snapshots
- Associating videos and snapshots with patient data
- Selection, display, review and comment annotation of images and videos
- Archiving of recorded image and video documentation

## Connection (Caliop con)

- Software connection to the hospital information system (HIS) and Picture Archiving and Communication System (PACS) for long-term archival
- Import of patient information from the hospital systems (HL7, DICOM Worklist)
- Export of images created during surgery using DICOM Storage



# User Test Design

A total of five studies were carried out with the “medical specialists” user group. The group was comprised of 26 clinical experts, including 21 OR nurses or surgical assistants, 2 nurse anesthetists, and 3 surgeons from 7 different hospitals, each with relevant work experience in the operating room. A total of 42 test sequences were performed to validate the system. To identify potential usability issues, participants were asked to perform a series of common tasks in a simulated OR environment without prior instruction on using the hardware and software.

The study took place at the Institut für Notfallmedizin und Medizinmanagement (Institute for Emergency Medicine and Management in Medicine (INM)) at Ludwig-Maximilian University in Munich and was supervised by M3i staff. The participants were observed by camera from three angles to allow subsequent review of the subjects' actions while performing the specified tasks.

## Study environment and set-up

The study environment consisted of a fully equipped operating room with an integrated EIZO Surgical Panel outfitted with the Caliop software from the CuratOR product portfolio. The usability study and follow-up interview of the participants took place in separate rooms.

Each iteration of the study consisted of the following steps:

- Run the test sequence: Test of the user-friendliness and intuitiveness of the software
- Survey of occupational tasks and participant knowledge
- Evaluate usability along the System Usability Scale with a questionnaire
- Review the subjectively experienced workload on the participant (NASA-TLX)
- Openly structured expert interview

## Participant profiles

The goal of the first test was to assemble a broad range of opinions in order to incorporate the widest spectrum of expert views in the product development process. This established areas of focus for the follow-up testing, such as differences in the technical affinity of the participants or their behavior in stressful situations as fabricated during a targeted simulation of an OR workflow.

## Data collection

The test sequences were divided into multiple tasks. After performing the test sequences, the System Usability Score (SUS) of the Caliop software was assessed. In the final test, the NASA-TLX Index was also measured. For the expert interviews, a general questionnaire was developed and adapted to the specifications of the testing performed in each iteration of the study. All test sessions were recorded on video to allow verification of the data in a subsequent review.

# Summary of Test Results

The study evaluated a range of aspects of the Caliop OR software as well as the seamless, manufacturer-independent interaction with connected components in the OR. This section presents the most important results and findings from the study.

## „EXPERIENCE THE FUTURE OF OPERATING ROOM TECHNOLOGY“

The ever increasing number of medical imaging and information systems represents a major challenge for some users in the OR. Proper operation of the equipment must be ensured, even in stressful situations. Different user interfaces and incompatible systems can force cumbersome workflows onto the OR team. With CuratOR Caliop, EIZO has developed an integration software suite that can centralize essential OR workflows on a single platform such as the Surgical Panel.

The first study offered useful findings about the intuitive operation of the software and integration with the Surgical Panel. Ambiguous terms such as "remove" and "delete" were reformulated to help users better understand the associated actions, and check boxes were introduced to simplify selection processes. An early version of the layout icon was replaced to prevent confusion with other system functions. Thanks to the participants feedback, the date of birth was added to the main display to ensure a quicker and more reliable identification of the current patient. The method for selecting records for archival was also redesigned for easier use based on the observation of the user interactions.





# Summary of tests

The study evaluation comprised three test iterations. Each was conducted with participants from the OR nurse and surgical assistants user group. Surgeons with specialties in oral and maxillofacial surgery, general surgery, abdominal surgery and transplantation were also included in the first study. Each test had a different focus.



## Overview of the various tests

The three tests were conducted in 2018 and 2019. A total of 26 different clinical users from 7 different hospitals participated in the tests.

Test	Time period	Number of participants	Test specification
1	December 2018	10	Initial criteria with focus on gathering information for future tests
2	May 2019	8	Focus on the technical affinity of the subjects
3	October – November 2019	15 (study was repeated with 5 participants)	<ul style="list-style-type: none"><li>• Simulation of a real OR situation</li><li>• System operation under stress</li><li>• Evaluate the learning curve</li></ul>

# TASK ASSIGNMENT

## Overview of tasks to be performed



1. ID 8 System start up  
Power on the Surgical Panel and start the software
2. ID 2 Patient administration  
Select, create, and activate patients
3. ID 10 WHO checklist  
Perform time-out in accordance with WHO checklist
4. ID 6 Verify patient  
Verify patient
5. ID 3/ ID 1 Video Management  
Assign images from various sources to connected targets and record videos and snapshots for procedure documentation
6. ID 7 Time and Stopwatch  
Activate and monitor timer and stopwatch
7. ID 5 / 4 Archiving  
Open, select, annotate and archive videos and snapshots created during task 5
8. ID 9 System shutdown  
Close software and power down the Surgical Panel

# Evaluation of tasks

**Studies show that a lack of user-friendliness and poor user interface design negatively impact medical devices [1]. The Caliop software suite from EIZO’s CuratOR portfolio helps physicians and medical experts concentrate on the clinical intervention, without having to think about the operation of an ever growing number of complex systems.**

A number of system use cases were evaluated prior to conducting the study, and multiple scenarios were developed to test these use cases. The tests included the tasks described above, and subjects were expected to perform them without prior instruction on the use of the system.

In the first two tests, all eight scenarios were validated. In the third test, four selected use cases of particular importance for the further development of the system were reevaluated in a simulated stress situation. After performing each task, subjects were asked to judge the level of difficulty on a scale of 1 to 5, with 1 signifying "simple" and 5 indicating "difficult".

	1st Test	2nd Test	3rd Test
System start up	1.5	1.1	–
Patient administration	2	2	1.7*
WHO checklist	2	1.6	–
Verify patient	1.9	1.1	–
Video Management	2.3	2.4	2.3
Timer and stopwatch	1.5	1.6	1.3
Archiving	2.5	2.9	2.7
System shutdown	1	1.4	–

\* The "Patient Administration" and "Verify Patient " use cases were tested in a joint scenario during the third test. The result of the participant’s evaluation of these tasks is included in the "Patient Administration" use case.

The subjects performed the tasks without prior instruction on the use of the system, and overall they deemed the interactions to be easy to moderately difficult. Almost all test participants judged the "System start up", "Patient administration", "WHO checklist", "Verify patient", "Timer and stopwatch" and "System shutdown" tasks as easy to moderately easy. "Archiving" and "Video Management" were deemed the most difficult tasks.

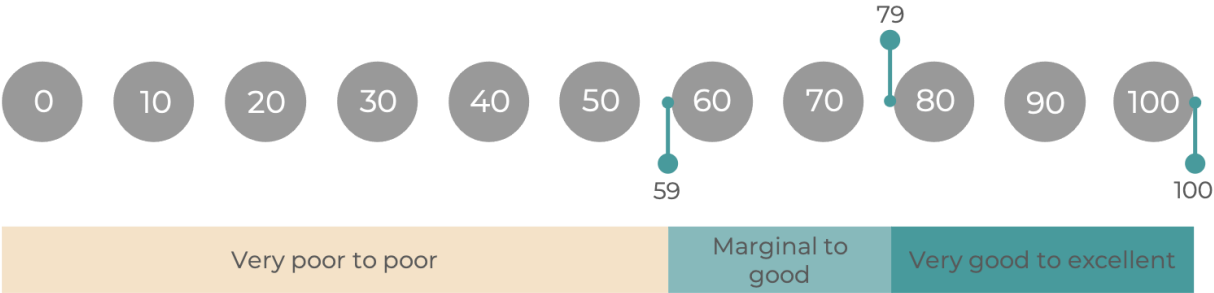
[1] Gurses A, Ozok AA, Pronovost PJ. Time to accelerate integration of human factors and ergonomics in patient safety. BMJ Qual Saf. 2012;21:347–51.



# Evaluation of usability

With a value of 81.1, the Caliop OR software rated above average on the System Usability Scale (SUS).

## System Usability Scale



The System Usability Scale (SUS) is a means of evaluating user-friendliness. It is an independent, scientifically tested scale that assesses the usability of technological systems, including household devices, high-tech systems and medical technology solutions. Users react to 10 statements on a scale of 1 to 5 (Likert scale), where 1 means "Disagree completely" and 5 means "Agree completely". The results are then converted to an overall score between 0 and 100.

After performing the test sequence, the system SUS score was calculated for all three tests. The following questions were asked to determine the SUS score:

The SUS is a 10-item questionnaire with 5 answer options for each question.

- ☐ I think I would enjoy using this system more frequently.
- ☐ I found the system to be unnecessarily complex.
- ☐ I found the system easy to use.
- ☐ I believe I would need the help of a technically adept person to be able to use this system.
- ☐ I found the various functions were well integrated into the system.
- ☐ I think the system has too many inconsistencies.
- ☐ I can imagine that most people learn to use this system quickly.
- ☐ I found the system awkward to use.
- ☐ I feel very confident using the system.
- ☐ I had to learn a lot before I could begin to use the system.

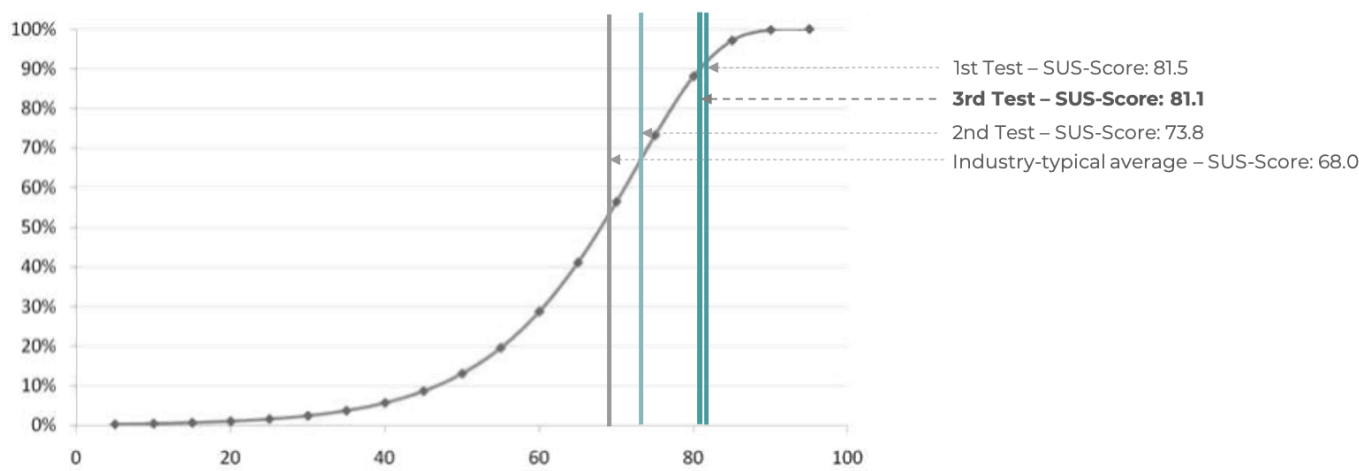
### SUS score: Comparison of all tests

After each test session, the participants were asked to respond to the SUS statements. In all three tests, the Caliop system achieved a significantly higher SUS score than the industry average of 68 [2].

Item	1st Test	2nd Test	3rd Test
1	3.1	2.4	3.4
2	3.1	3.0	3.3
3	3.0	2.9	2.6
4	3.6	2.9	3.6
5	3.5	2.8	3.4
6	3.1	3.1	3.1
7	3.8	3.3	3.3
8	3.4	3.4	3.7
9	2.5	2.4	2.6
10	3.5	3.5	3.6
Total	81.5	73.8	81.1*

\* Revised SUS score (14 of 15 subjects), one value was determined to be an outlier (> 2.5 standard deviations) and therefore was not included in the evaluation

The Caliop system received an SUS score of 81.5 in the first test. This represents very good usability. The second test was targeted specifically to be conducted with subjects to have low technical affinity. As expected, the SUS score dropped to 73.8. The revised SUS score of 81.1 determined in the third study is the most significant because of its ecological validity and higher number of subjects, and is therefore included in the evaluation as the primary SUS score.



[2] Jeff Sauro. A Practical Guide to the System Usability Scale. CreateSpace Independent Publishing Platform, 2011

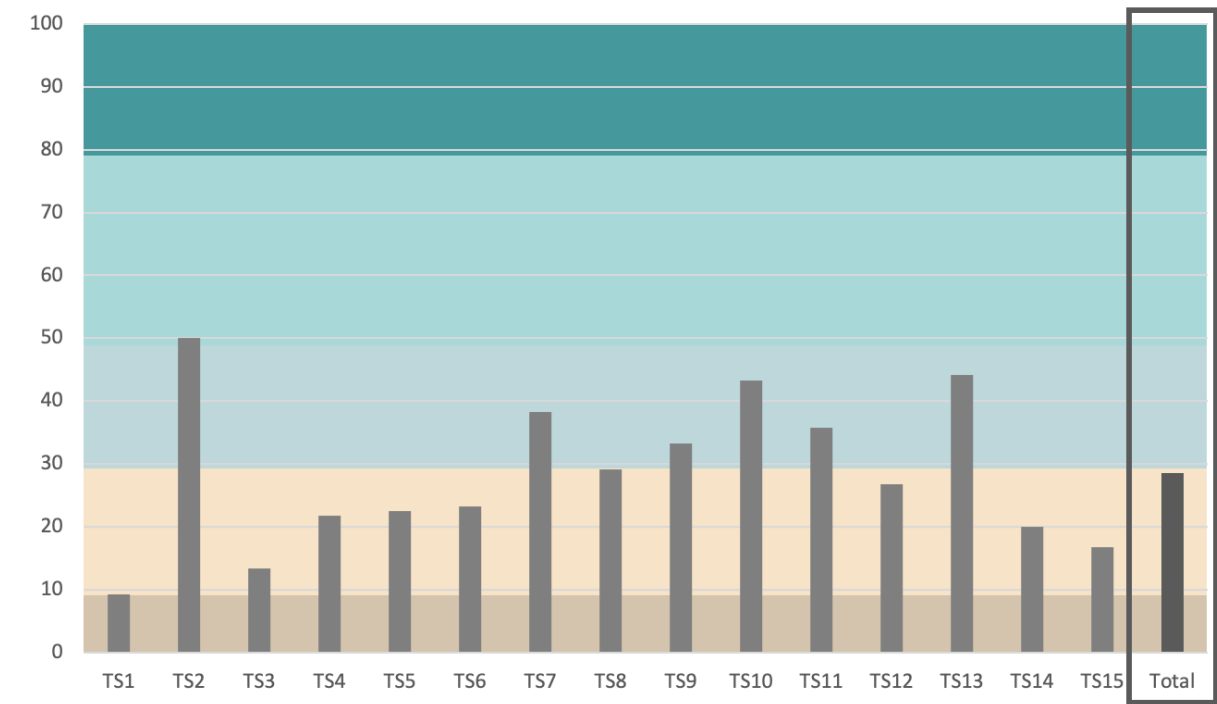
# Measuring the subjective workload of subjects in a fabricated stressful situation

## NASA-TLX

The NASA Task Load Index (NASA-TLX) by Hard & Staveland measures the subjectively experienced workload using a multidimensional construct and serves as a standard procedure for determining the mental strain on a person. Six dimensions are assessed: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. To enable an evaluation, the NASA-TLX is divided into 5 levels of workload. The scale is from 0 to 100 [3]:

Low subjective workload	Moderate subjective workload	Somewhat high subjective workload	High subjective workload	Very high subjective workload
0-9	10-29	30-49	50-79	80-100

A NASA-TLX index of 28.5 was determined as the median across all participants. During this test, the participants were intentionally subjected to an increased workload, time pressure and interruptions during parallel activities. The evaluation of the subjective workload shows that overall the subjects felt they were under moderate workload when interacting with the Caliop software and the Surgical Panel during the test sequences.



[3] Barbara-Ulrike Groß, Development of Degrees of Difficulty in a Test Environment to be Developed (Entwicklung von Schwierigkeitsgraden in einer zu entwickelnden Versuchsumgebung), 2004



# Evaluation of the Learning Curve

The following details the results of the system learning curve evaluation. The testing was repeated with a third of the subjects after approximately one week. The tasks were left unchanged in order to minimize potential confounding factors.

## Change in task evaluations

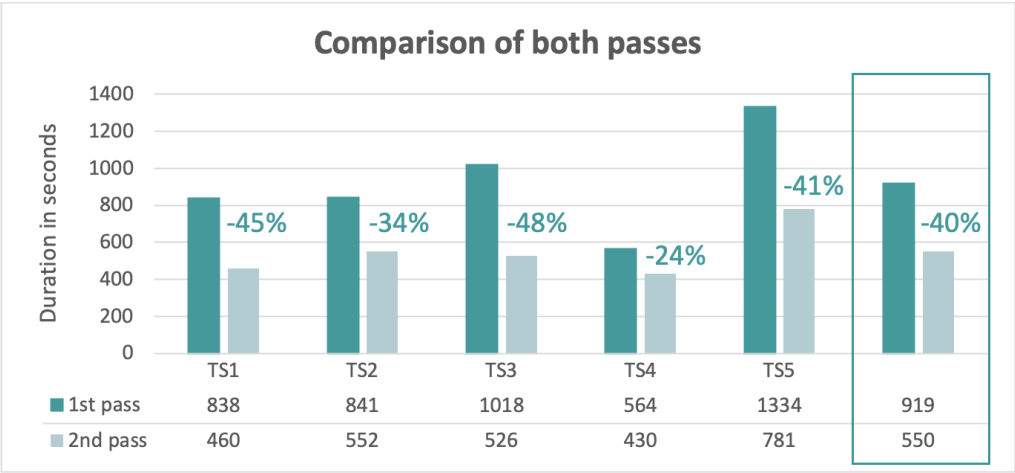
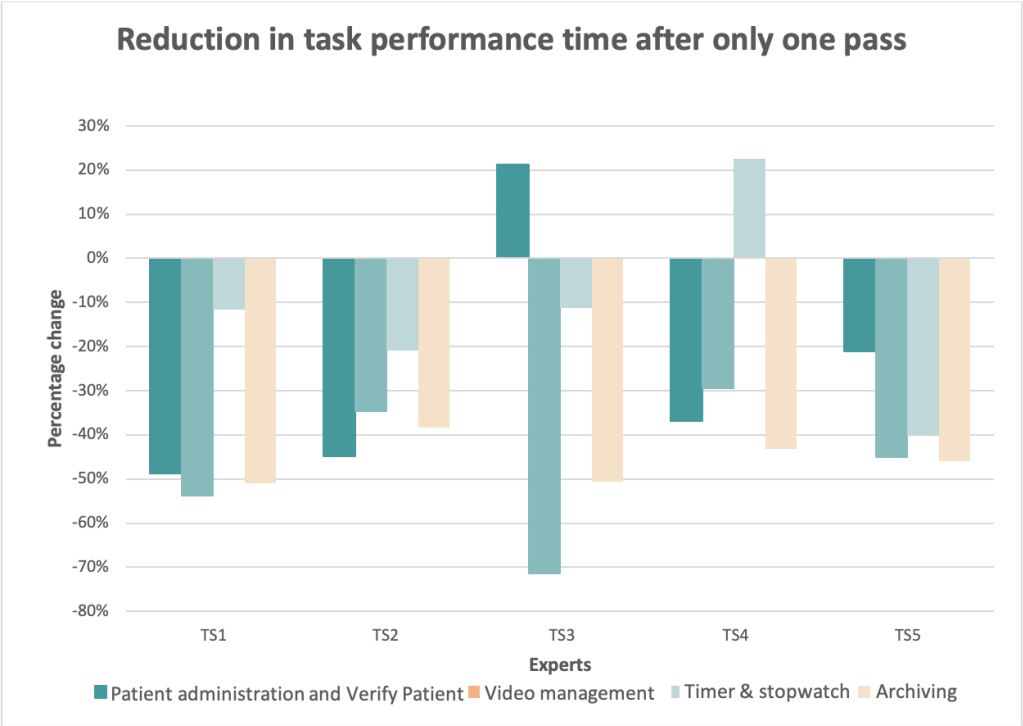
To start, the change of key quantitative data used to evaluate the first session was tabulated and analyzed. The values from the five participants first session (1st pass) were compared with the values of the second session (2nd pass), and the changes between the two passes were identified.

During the two passes the subjects received no instruction on the use of the system and had to learn how to work with the system intuitively. The Video Management and Archiving tasks judged more difficult in the first pass were perceived as easier by all subjects in the second pass.

	Difficulty rating								Total	
	Patient administration & Verify patient		Video Management		Timer & stopwatch		Archiving			
Pass	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
TS1	1	1	2	1	1	1	2	1	1.5	1
TS2	1	1	2	1	1	1	2	1	1.5	1
TS3	2	1	3	1	1	1	2	1	2	1
TS4	1	1	2	1	3	2	2	1	2	1.25
TS5	1	2	2	2	2	1	3	2	2	1.75
									1.8	1.2

On average, the difficulty rating dropped by a third from 1.8 (during the first session) to 1.2 (during the second session). This shows that subjects found it easier to operate the system during their second interaction and retained the understanding they derived during the first pass.

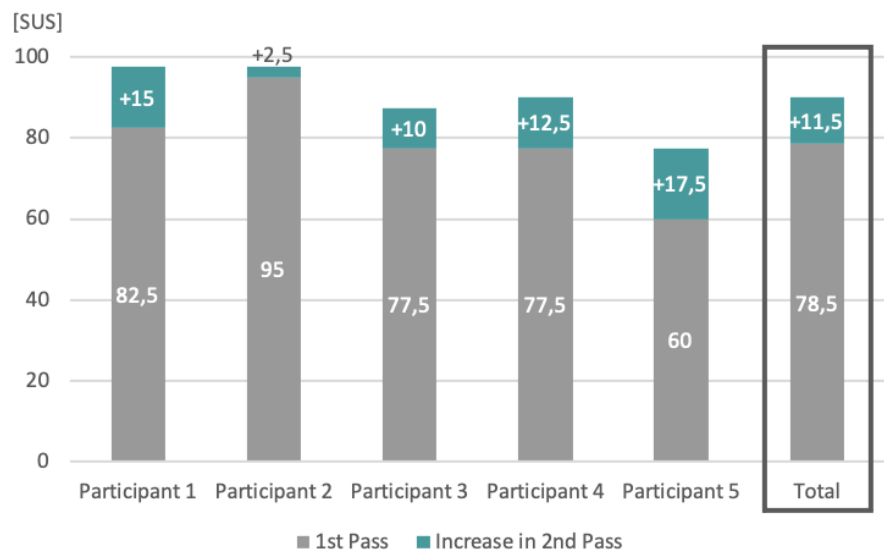
Change in task performance time



On average, the total task performance time dropped by 40.2 percent, from 919 seconds (during the first session) to 549.8 seconds (during the second session). The reduction in the key metrics shows that the test subjects were more proficient using the system during the second pass and indicates that a learning effect had occurred and the users retained their intuitive understanding of the system from the first pass. In many cases, the five subjects had a notable drop in percentage values in both their evaluation of task difficulty and in their task performance time. Where values were already low in the first pass, no change or a slight increase were noted on an individual basis.

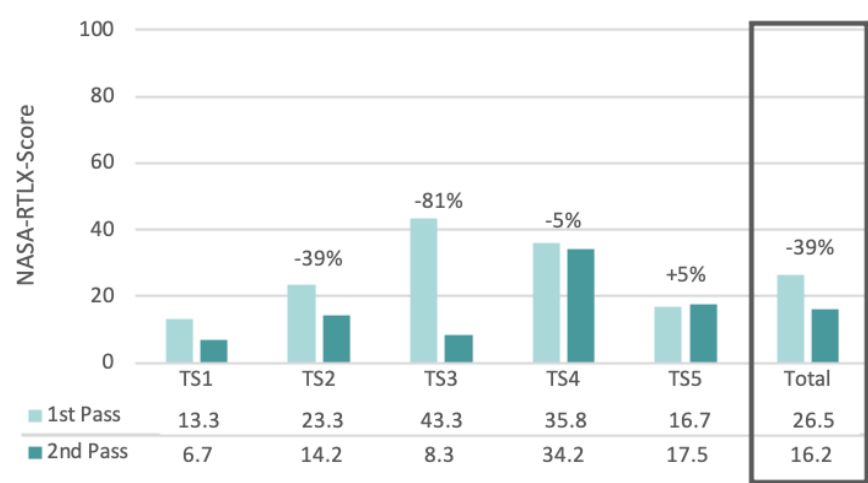
### Change in the evaluation of usability

The second pass showed an improvement in the SUS score for all five participants. Across all test subjects, the SUS score improved by 11.5 points, from 78.5 to 90, which represents a 14.7% increase. As such, the system exhibits very good to excellent usability.



### Change in subjective workload

The questionnaire was used to determine the raw TLX (RTLX) index, which is a form of the NASA task load index without the subsequent weighting of the six dimensions.



A reduction in the subjective workload was noted in the second session for four of the five participants. Across all test subjects, the NASA-RTLX score decreased by 10.3 points, from 26.5 to 16.2, which represents a drop of 38.9 percent. Based on the interpretation guideline, the NASA-RTLX score of the second session indicates a moderate subjective workload.





## Summary

This study delivers reliable and objective data from three tests. By using proven methods such as the SUS score and NASA-TXL to test usability and workload, the collected data are easily comparable and deliver valuable findings about the perceptions of the intended user group. The results represent a positive validation of the envisioned benefits of the system:

- In all three tests, the tasks were largely evaluated as easy to perform.
- An SUS score of 81.1 was achieved, which is an above-average value and represents very good usability.
- Across all subjects, the NASA-TLX Index had a calculated value of 28.5. This represents a moderate subjective workload.
- The review of the learning curve showed improvement in all areas observed: the assessment of the degree of task difficulty dropped by 33%, the total task performance time fell by 40%, the SUS score increased by 15% and a drop of 39% was noted for the NASA-TLX index.

The overall results of the study confirm that the Caliop software and the Surgical Panel provide positive support for OR personnel by centralizing critical OR workflows on a single platform, resulting in considerable time savings. Although the Caliop software offers extensive functionality, all of the participating experts said the system was easy to use. This is critical for modern treatment environments, where different medical users and a large number of information and video sources are involved.



Performed and evaluated by



INDUSTRY-IN-CLINIC PLATFORM