Usability assessment of an intraoperative planning software

Usability Assessment of an Intraoperative Planning Software

Federico Sternini[1][a], Giuseppe Isu[2][b], Giada Iannizzi[2], Diego Manfrin[2], Noemi Stuppia[1,3], Federica Rusinà[1] and Alice Ravizza[1][c]

[1]*USE-ME-D srl, I3P Politecnico di Torino, C.so Castelfidardo 30/a, Torino, Italy*
[2]*Medics srl, via Avogadro 19, Torino, Italy*
[3]*Politecnico di Torino, C.so Duca degli Abruzzi 24, Torino, Italy*

Keywords: Usability, Learning Curve, Touchless.

Abstract: Usability is a crucial aspect of medical device safety. The brand-new European Regulation requires the manufacturer to assess the usability of the new medical devices. In this study, we evaluate the usability of a new medical device intended to assist the intraoperative planning with the visualization of 3d patient-specific organ models. The usability study started from the early stage of the device design and iterated through an early formative, completed with desk-based activities, late formative, completed with a focus group, and summative phase, that comprised a user test, and questionnaire filling. The identified usability issues are mitigated, the safety of the device user interface is confirmed and the training contents are defined and confirmed. Additional information regarding the user experience is collected and analyzed to identify further improvements of the device.

1 INTRODUCTION

Usability assessment of medical devices is becoming a widely diffused practice during device design. The diffusion of this practice is partly eased by the European regulatory framework for medical devices. Medical device regulation 2017/745 (European Parliament and of the Council, 2017) requires that risk evaluation includes the evaluation of risks and hazards related to human factors.

During the design of the medical device, object of this study, the methodology for the assessment of the human factors follows the relevant international standards. The international standards define a method designed to ensure a high-level quality of the medical device interface in terms of safety for both patients and operators. The method foresees an iterative workflow, that requires different steps to be completed. The first phases are so-called formative, which are used to define the interface design and to establish the details of the device design. In the later phases, the confirmation of the user interface safety (called “summative”) is completed.

2 MATERIALS

2.1 Device

The device assessed is a software as a medical device (SaMD) intended to aid the surgeon for the intraoperative planning thanks to the presentation of 3d reconstructed models of the patient-specific anatomy. Briefly, the models are realized as based on the radiological images of the patient (e.g. CT or MRI) through the segmentation of the 2D medical images. The obtained 3D models are then made available to the physician through a proprietary platform. In the platform, the physician can add notes, information, and custom requests to the model. Once the model is confirmed by the physician, it is made available in the device ICON, which accesses the platform. Thus, the model can be visualized in all its parts. The visualization is aided by a touchless user interface enabled by the LEAP MOTION sensor (Ultraleap, US), that tracks and identifies the hands of the users, without the need for additional sensors. The user can modify the visualization of the organ model

[a] https://orcid.org/0000-0002-5510-2296
[b] https://orcid.org/0000-0002-8440-607X
[c] https://orcid.org/0000-0003-2368-7258

Sternini, F., Isu, G., Iannizzi, G., Manfrin, D., Stuppia, N., Rusinà, F. and Ravizza, A.
Usability Assessment of an Intraoperative Planning Software.
DOI: 10.5220/0010252904830492
In *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2021) - Volume 5: HEALTHINF*, pages 483-492
ISBN: 978-989-758-490-9

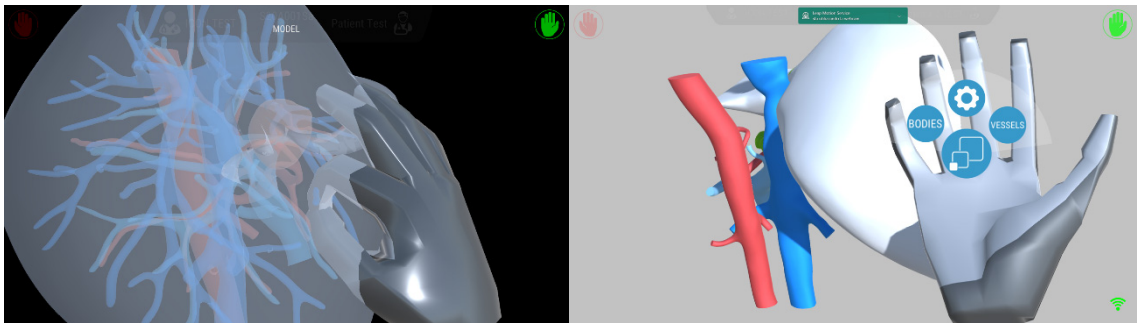

Figure 1: Examples of device interaction. On the left, the user is rotating the model, as can be seen by the “pinching” gesture. On the right, representation of menu opening.

in terms of positioning, zooming and orientation. In addition, the user can select the visibility of model specific parts selecting one of the three visibility statuses: solid, transparent, and hidden. The management of the visualization is completed with three main hand gestures and with interaction with a menu. The gestures are the following:

- Gesture for the model rotation. The user shall place the hand on the 3D model and subsequently pinch point finger and thumb together. Keeping the pinching, the user can move the hand in any direction in the space and the 3D model will start rotating following the hand rotation around the center of mass of the model.
- Gesture for the model panning. The user shall place the hand on the 3D model and subsequently close the hand in a fist. The user can translate the 3D model moving the hand in any direction in the space while keeping the fist.
- Gesture for the model scaling. The user shall place both hands open to surround the 3D model. Subsequently, moving the hands away from the 3D model the model will start scaling up; moving the hands towards the 3D model, the model will start scaling down.

The visibility status of all components of the organ model can be managed using a floating menu. The floating menu can be opened by the user rotating the hand palm up, and the menu will be positioned in the user's palm. The menu presents round buttons that represent the main categories of the model elements, i.e. bodies and vessels, and the button to access the setting section. Once one category is opened, the elements belonging to the selected category are presented and available for selection. Interaction with the buttons composing the menu is completed by pressing the circles.

3 METHODS

As suggested by the international standard IEC 62366 (International Electrotechnical Commission, 2015), the methodology for the usability assessment was structured in two phases: first a formative evaluation and then a summative evaluation.

3.1 Formative Evaluation

The formative evaluation is the phase intended to iterate the device design until a satisfactory quality level is reached. The formative evaluation of this device was designed in two separate phases. As the formative evaluation began during the early phase of the development of the device, the first phase was desk-based, while the second phase comprised the participation of real users as participants to a focus group.

In the first step, designers and usability experts used techniques considered appropriate to the design development stage in terms of outputs and resources needed (Ravizza et al., 2019). The team used a quick and dirty approach and used low resources techniques listed in the IEC 62366, such as brainstorming, FTA, cognitive walkthrough, and standard review. The outputs of the first phase included the definition of a set of primary operating functions, i.e. the functions that the user shall be able to complete to achieve the intended use, that have to be evaluated in the next phases. Additionally, this first phase had as output the definition of the position of the sensor and the screen to allow correct ergonomics of the user.

The following phase was the focus group. This technique was planned at this stage in the usability evaluation to confirm the outputs of the previous stage and to identify possible additional issues thanks to the analysis of the end-user perspective. The focus group was organized during the Covid-19 pandemic and

therefore required the moderator to assist only one participant for each session; the consensus statement of the participants was obtained by virtual meetings. Each focus group session was then structured into 4 brief sections, for a duration of a maximum of 90 minutes. The first section included a brief training for the device use. This section lasted a maximum of 30 minutes and provided the minimum information required to use correctly the device to the users. The training session is designed based on the previous outputs and is designed to be consistent with the training that will be provided to actual users. The focus group training was used as a basis for the future commercial training required after the distribution of ICON to the customers. After the training session, the users were invited to complete a set of tasks with the device. After the completion of the tasks, the users were asked to provide an evaluation of the primary operating functions and to provide information regarding some crucial aspects with a closed-ended questionnaire built on the base of a 5 point scale. The scale is designed to range from the value zero, which is associated with the absence of usability problems, to a value equal to four, which represents the presence of usability problems that could impact patient health. Finally, after the completion of the questionnaire, the users were invited to a discussion with the designers and the usability experts to find additional usability problems and to propose any suggestion for the user interface.

3.2 Summative Evaluation

The summative evaluation is the last phase of the usability evaluation and is intended to confirm the usability of the medical device. Therefore, the device involved in the study shall be consistent with the final version of the medical device and shall present all the features of the medical device.

After the completion of the formative phase, the device user interface received the following modifications, that impacted the primary operating functions and the structure of the summative evaluation: the positioning of the sensor, of the screen, and the parameters of the virtual view, are set by the manufacturer, and a tutorial section is included in the device to allow the users to familiarize with the gestures and the menu structure. The tutorial section contents were obtained from the training contents identified during the formative desk-based phase and the analysis of the issues presented by the users during the focus group.

The summative evaluation of the medical device involved final users in sessions of simulated use of the device. The simulations were completed in a setting intended to represent the real setting of the medical device inside the operating room. Therefore, the simulated use setting included the provision of a surgical column, consistent with the column that will be provided by the manufacturer to users, equipped with a medical-grade workstation for the software proper execution, a medical-grade screen, and a flexible arm for the sensor placement. The column was placed on one side of a table covered with cloths intended to mimic the sterile drapes usually placed on the patient during surgical procedures. Also, as the device allows for the visualization of the virtual model combined with a video stream collected from external video sources, a simulation of a patient undergoing a laparoscopic procedure was realized by a closed box containing the tip of the video laparoscope and a 3d printed model of the liver. The model was the physical print of the same model presented to the user inside the medical device.

The user test is structured in different phases.

3.2.1 Training

Training: the design team presents the medical device to the user, explaining all the relevant information for the device use. This information included the gestures required for the device interaction and the tips intended to ease the first use of the medical device.

3.2.2 Task Analysis

Task analysis: the moderatos asked the user to complete some complex actions while observing the device use and annotating the performance of the user for each task. The moderators classified each task completed by the user in one of the following 4 classes:

- Ok: the user completed the task correctly.
- Ue: use error. It represents any task that the user was unable to complete, that was completed without awareness of its meaning, that was completed by mistake, or that required intervention of moderators.
- Te: technical error, represents the cases when the device presented some technical issue that did not allow the user to complete the task.
- C: critical, represents particular cases of use errors that can be associated with an impact on patient health

3.2.3 Heuristic Evaluation

After the completion of the simulated tasks, all users were asked to compile a questionnaire for the heuristic evaluation of the device. The heuristic analysis is an inspective technique intended to identify the elements that violate the usability

heuristics (i.e. identify usability problems in the user interface). After the identification of the violations, a score is assigned to assess the severity of the violation (Zhang et al., 2003).

The designer and the usability experts designed the questions, analyzing the heuristic principles proposed by Zhang et al (Zhang et al., 2003) and proposing a set of questions designed to fit the user interface features of the device under assessment. The presentation of closed-ended questionnaires allowed for the evaluation of the severity of heuristic violations, even if the user is not an expert in this technique. The user could answer each question using the same scale proposed for the questionnaire proposed during the focus group to maintain the consistency of the test methods across the different stages of the usability evaluation.

3.2.4 Primary Operating Functions and Risk Questionnaire

Later the user was asked to fill a questionnaire consistent with the one proposed during the focus group, intended to exploit the crucial aspects of the user interface of the device and its primary operating functions. As the heuristic evaluation, the scale for the answers is the same proposed during the focus group.

3.2.5 UEQ Questionnaire

After the completion of the two first questionnaires, the user was asked to complete a third questionnaire, that is not relevant for the evaluation of the risk profile of the device, but that is intended to describe the overall usability of the user interface. The UEQ questionnaire is a standardized questionnaire and presents a set of couples of terms, and the user has to select the evaluation of the device for each couple of terms, positioning the device evaluation in the scale described by the terms (Laugwitz et al., 2008).

3.2.6 UEQ Questionnaire Stereoscopy

After the completion of the three questionnaires, the moderators asked the participants to try a different visualization mode provided by the medical device. This mode is designed to allow the use of the model in stereoscopic screens and displays. The simulation was completed with a virtual reality visor. As the setting of the simulated setting was not representative of the real medical device use, the tasks are not evaluated as in the previous stages, but an additional questionnaire was proposed to the participants, asking them to fill the questionnaire considering the stereoscopic visualization only.

4 RESULTS

4.1 Formative Evaluation

4.1.1 Desk-based Phase

During the desk-based activity, the usability and design team identified the positioning for the LEAPMOTION sensor and the screen that allows the user to have a comfortable organ model visualization. The frontal positioning of the screen and sensor allows the user to have a visualization consistent with the placement of the hands.

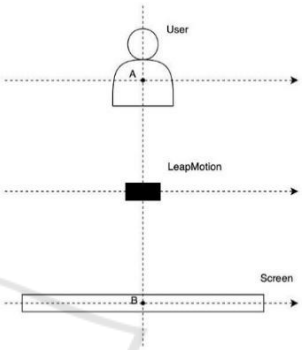

Figure 2: Scheme of the positioning of the user, sensor, and screen.

After the identification of the positioning, the primary operating functions are identified, here listed in Table 1.

Table 1: Primary operating functions.

Primary operating functions	Interface testable technical requirements
Choose the case	The user shall be able to select the proper case use (organ model) for the surgery
Set up of the operator against the virtual view	The information provided by the system shall allow to set up appropriately the user position against the virtual view
Handling the organ model	The hand gestures shall allow to manipulate the organ model in an intuitive manner
Management of the parts belonging to the organ model	The hand gestures and the menu setting shall allow to isolate and change the transparency of the parts belonging to the organ model in an intuitive manner
Management of the scene background	The user shall be able to switch the scene background in an intuitive and simple manner

4.1.2 Focus Group

The focus group was completed with the participation of four users, in line with standard recommendation, which suggests at least 4 participants and a maximum of 8 participants (International Electrotechnical Commission, 2016). The four users were all surgeons, two orthopedics, and two thoracic surgeons. During the task completion, the following issues were identified by users and moderators:

- The gesture required to zoom the model is not so intuitive.
- The hands should be placed in the sensing volume completing a predefined movement that allows easy identification of the hands (half-moon shape trajectory).
- The users had some troubles when trying to pinch and rotate the model.
- Some users found the position of the menu opening uncomfortable and would have preferred a position that does not require taking the hand backward to select the tiles
- Some users had some difficulties to visualize and read the menu elements due to the transparency of the menu overlaid to the model.
- One user could not use the gestures because kept the second hand in the sensing volume.

After the completion of the tasks, the users compiled the questionnaire, results did not present any value higher than two, which represents that the device made the user nervous. We recall that the higher is the score, the worse is the usability problem. In particular, the highest score was obtained by one single operator on the question regarding the possibility to manage the transparency of the model components. The details of the questionnaire answers are presented in Figure 3.

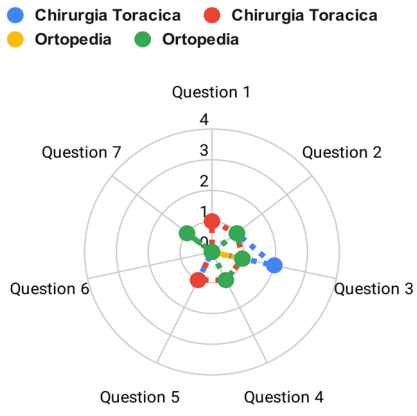

Figure 3: Details of focus group questionnaire results.

4.2 Summative Evaluation

The participants in the user test were 16. The number of participants is appropriate for the task of user interface safety confirmation, as 15 users are considered the minimum practical number for usability validation purposes (Center for Devices and Radiological Health, 2016). All of them were professional users. Five of them were urologists, seven were orthopedics, two were generic surgeons, one was an emergency surgeon and one was a thoracic surgeon, representing all the specialties that the manufacturer can provide with 3D patient-specific organ models.

4.2.1 Task Analysis

15 out of 16 users completed the list of requested tasks. The users completed correctly 60% of the tasks, while 20% of the tasks were classified as use errors. The remaining 18,59% of tasks were not performed by the users and technical errors occurred in 1,41% of the tasks.

The performance of the users could be divided into three macro tasks: first, the users are asked to complete the tutorial section of the device, then the users are asked to interact with the medical device while visualizing a hip arthroplasty model, and finally, the user interacted with a liver model. Considering the division in macro tasks, during the device use the user performed reducing the prevalence of use errors and increasing the number of not completed tasks, while the prevalence of the correctly completed tasks remained quite stable for all the phases of the test.

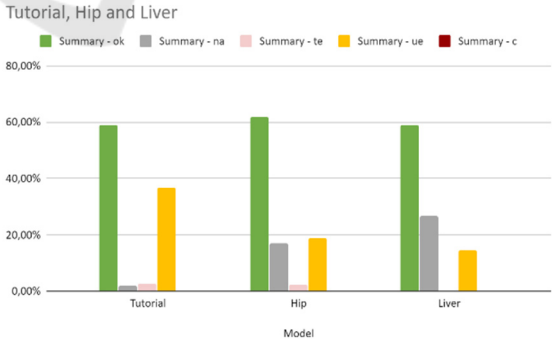

Figure 4: Performance of users divided per macro tasks.

4.2.2 Heuristic Analysis

All the participants filled the questionnaire for the heuristic analysis of the device. The users were never assigned a score higher than 2, which corresponds to

a violation of the heuristic principle that made the user nervous during the device use. Therefore, the user never answered that the user was impossible to use or that the device use could lead to an impact on patient health. The details of the aggregated answers are presented in Figure 5.

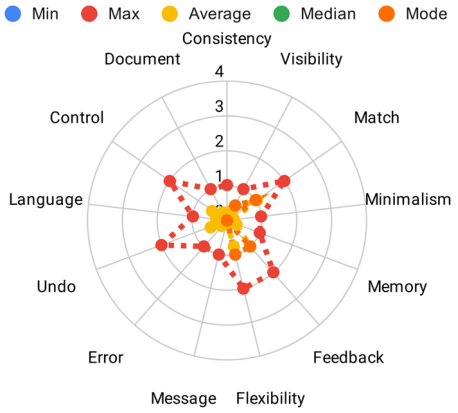

Figure 5: Minimum, maximum, average, median, and mode value of the scores assigned by the users to each relevant heuristic principle.

4.2.3 Primary Operating Functions and Risk Questionnaire

All the participants filed the questionnaire regarding the primary operating functions and the specific questions regarding the risks of the device. The primary operating functions were modified from the previous iteration of the usability testing due to the modification of the device. Therefore, the primary operating functions were defined as follows:

- Completion of the tutorial
- Choose the case
- Handling the organ model
- Management of the parts belonging to the organ model
- Management of the scene background

Also, three questions related to the completeness of the user interface, the color-coding, and the clarity of the notifications are asked. All participants but one assigned scores lower than two (device use made me nervous) to all the primary operating functions and situations related to the main risks associated with the device. The primary operating function that received a score equal to two is the one associated with the management of the parts belonging to the organ model. Details of the questionnaire results are presented in Figure 6.

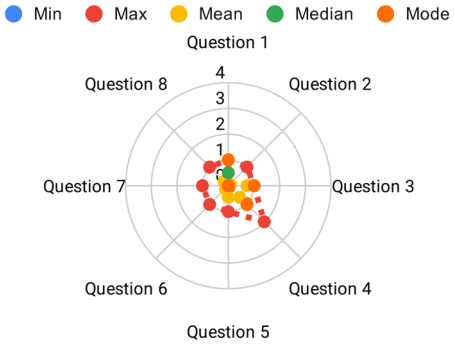

Figure 6: Minimum, maximum, average, median, and mode value of the scores assigned by the users to the five primary operating functions and the three risk-related questions.

4.2.4 UEQ Questionnaire

All participants filled the UEQ questionnaire. The questionnaire results were evaluated according to the dedicated data analysis tool, the results are cleaned removing the inconsistencies of the answers provided by the users, and the results detailed in Table 2 in the usability areas of the device are obtained (User Experience Questionnaire (UEQ), n.d.). The cleaning of data was completed by removing all the questionnaire data related to users that presented at least 2 inconsistencies among the answers, as it may be associated with low attention during the questionnaire filing.

Table 2: UEQ scales results.

UEQ Scales (Mean and Variance)		
Attractiveness	2.146	0.90
Perspicuity	1.750	0.98
Efficiency	1.917	1.40
Dependability	1.984	0.75
Stimulation	2.391	0.68
Novelty	2.484	0.57

4.2.5 UEQ Questionnaire Stereoscopy

Twelve out of 16 participants filed the UEQ questionnaire for the stereoscopy evaluation. The questionnaire results are evaluated as consistently with the other UEQ questionnaire as per the methodology proposed with the questionnaire (User Experience Questionnaire (UEQ), n.d.). The results are cleaned with the same criteria used for the other UEQ questionnaire. The results of the cleaned UEQ questionnaire are presented in Table 3.

Table 3: UEQ scales results for the stereoscopic visualization.

UEQ Scales (Mean and Variance)		
Attractiveness	2.403	1.02
Perspicuity	2.000	1.45
Efficiency	2.023	2.12
Dependability	1.977	0.72
Stimulation	2.477	1.43
Novelty	2.614	0.63

5 DISCUSSION

The study allowed the team to collect many data regarding the user interaction with the device, enabling the definition of improvements of the user interface and to define the device safety. During the formative stage, the usability issues identified during the focus group were analyzed and mitigated with different techniques.

The first methodology was the provision of adequate training to the user before the device use. In particular, the training focused on the position of the hands and a clear explanation of the gestures. The users had difficulties when completing the movements associated with the modification of the zoom of the model and the rotation of the model. In both cases, the clear explanation of the gestures with the provision of examples completed by the moderators allowed the users to improve their user experience and complete the tasks correctly. Therefore, the designers decided to introduce the tutorial section, intended to make the user practice with the gestures and have an easier interaction with the device. Other issues associated with the menu were resolved with a modification of the user interface, adding a back panel to the menu tiles presentation, reducing the visibility problems, and allowing the users to move the menu once opened, and to place it in a more comfortable position.

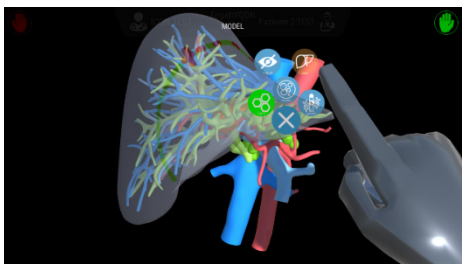

Figure 7: Menu visualization of the device version presented during the formative evaluation.

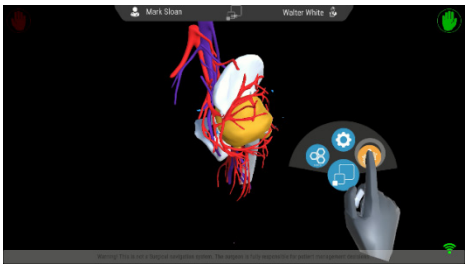

Figure 8: Menu visualization of the device version used during the summative evaluation.

The new functionalities are included in the version tested for the summative evaluation, and the tutorial section became an integral part of the simulated use testing.

The results of simulated use showed that the device cannot lead to risks for the patients, as the moderators did not classify any action as a critical error. Also, the technical errors were very few and led to a complete stop in the device use only within the tutorial section, which is not a medical module of the software and not intended to be used during the intraoperative planning. The results also showed that the percentage of use errors decreased rapidly during the device use, suggesting that the users can learn the correct use of the device very quickly during the device use, producing a steep learning curve when compared to the curves associated with surgical procedures (Hopper et al., 2007). Further observations can be completed by removing from the analysis user #13, which completed only the first part of the simulated use, completing only 13 out of 40 tasks, and removing the not completed tasks. The number of not completed tasks is affected by a set of 6 tasks that were misleading to the users. These tasks are the ones associated with the possibility to hide or make transparent all the components of the model belonging to a specific category. Many users completed the task hide or made visible the elements of the entire category (e.g. veins, arteries, etc.) without using direct command buttons, but by performing more commands than required. Therefore, these tasks were recorded as not completed, but not as use error, because the goal of the tasks was correctly reached.

When removing both these data from the analysis, the improvement of the user performance during the device use is more evident and recognizable. The percentage of correctly completed tasks increased at each macro task, ranging from the minimum of the first phase of use equal to 62,14% to the maximum reached in the last phase of the device use equal to 80,40%, while the percentage of the use error

decreased ranging from the maximum of the first microtask, equal to 34,95% to the minimum reached in the last phase of use equal to 19,60%.

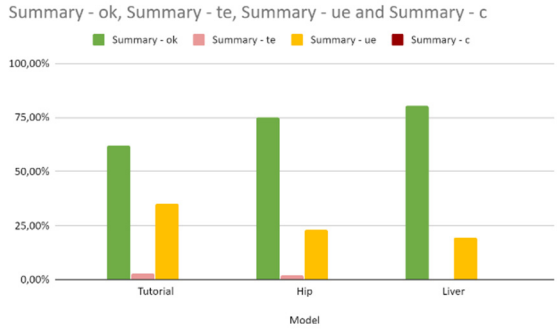

Figure 9: Analysis of task performance after removal of user #13 and the removal of the not completed tasks.

Considering the strict protocol of the test, the results are very promising. As a test rule, the moderators were not allowed to help the users during the test, neither to answer the user questions. In the case of help provision from the moderators to the users, the task was recorded as a use error. Therefore, the data show that the user required less training or help during the session and that they could remember the information required for the correct use of the device with a brief training session and few questions during the first device use.

Nevertheless, during the questionnaire filling phase, the users pointed out that they had difficulties during the management of the model and the use of gestures. In fact, in the heuristic questionnaire, the highest scores are associated with the following heuristic principles:

- Match: relevant for the consistency between the gestures of the hand and the commands received from the software. The mismatch could be caused by many factors, but the most prominent one is the lack of training. All users tried to complete actions with improper hand gestures.
- Feedback: principle applied to feedback provided by the menu interactions. The menu is the most difficult part of the software to use, as it requires the user to have confidence with the correspondence between actual hand position and virtual hand position. Also, the color code of the menu is intended to ease the comprehension of the menu parts status, but at the moment of the study was graphically presented to the user with a use example and as part of the brief training received by the users. No legends were presented to the user in any part of the user interface.
- Flexibility: the main issue that users identified is that the device requires attention and can be tiring to use during surgery. The users completed an intensive test that lasted from one to two hours, while during surgery the device will be used for a few minutes.
- Undo: The main difficulties are tightly related to the navigation of the menu, as the visualization status is controlled by the menu.
- Control: this heuristic principle is tightly related to the menu navigation, as the control of each model component is completed through the menu. Therefore, the difficulties identified by the users could be related to difficulties in the menu navigation.

The difficulties associated with the menu are directly related to the interaction that the user has to complete to modify the visualization status of the model. The user opens the menu rotating up the palm, then selects the tiles of the menu as if they were physical buttons, so the user had to press the tiles and then retract the finger from the selection. While the interaction is intuitive and does not require training to understand the movement, it requires that the user is aware of the position of the hand in the real and virtual representation of the space. The brief duration of the training phase and the short duration of the test could lead to the difficulties of the user to have fine control of the movements of the hand in the virtual space, and then, as the menu is the part of the interface

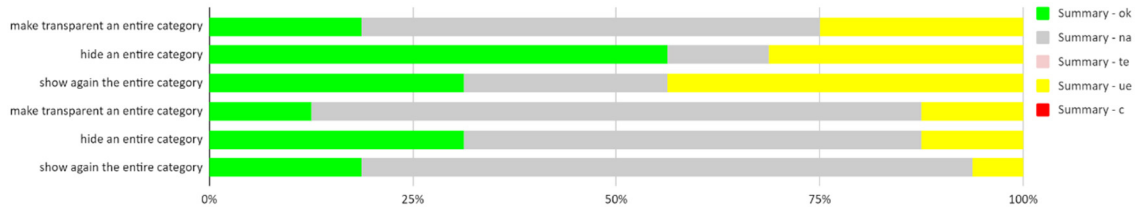

Figure 10: Representation of the performance of the users in the misleading tasks: the percentage of not completed tasks is high due to the possibility of completing the same task with different methodologies.

that requires the most precise interactions, the difficulties in the menu interactions. In these regards, during the design phase, the developers included the possibility to interact with the menu also with the traditional mouse/touchpad interaction, to help the user in the first uses of the device difficulties. During the summative tests, the users were not deliberately instructed about the possibility to use the mouse, to strictly evaluate the usability of the device by the innovative LEAP MOTION controller. This aspect enforces the positive outcome of the test considering that the possibility to have a well-known backup solution in case of difficulties constitutes a distress-relief and facilitation to accelerate the learning curve.

The same observations could be done on the results of the UEQ questionnaire. The questionnaire is designed to evaluate the aspects of the user experience. In all of these aspects, the device is considered very good, as the mean score is always higher than 1,6 which is twice the value considered for a good result (0,8). Even if the sample of users is quite small, the confidence intervals of the scores are always higher than 0,8.

Even if the observation from the task analysis led the moderator to the conclusion that the learning curve of the users when using this device is very steep, the users found perspicuity, which is the aspect that describes how easy is for the user to learn how to use the device, the worse usability aspect of the device. Nevertheless, even this aspect, have a positive score on average.

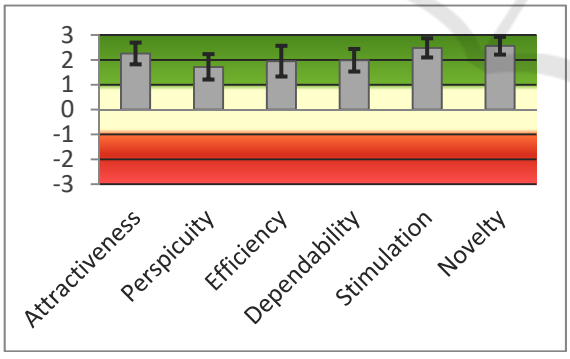

Figure 11: Scores of the UEQ Questionnaire and the associated confidence intervals.

Similar results are obtained from the evaluation of the UEQ questionnaire proposed to the user regarding the stereoscopic visualization only. Between the two questionnaires, no statistically significant differences (p=0,05) in the scores of the usability aspects are evidenced. Nevertheless, when comparing the results of the two questionnaires, the greatest difference is perceived in the perspicuity aspect. The entity of the difference may be justified by the possibility of the users to perceive more easily the hand position when interacting with the menu, thanks to the capability of presenting the third dimension provided by the stereoscopic visualization.

Table 4: T-Test for the difference of the UEQ scores between the general and stereoscopic visualization.

Attractiveness	0.7007	No Significant Difference
Perspicuity	0.5346	No Significant Difference
Efficiency	0.8863	No Significant Difference
Dependability	0.9861	No Significant Difference
Stimulation	0.9884	No Significant Difference
Novelty	0.8773	No Significant Difference

6 CONCLUSIONS

The study allowed to identify the issues of the user interface of the device at the design stage and allowed the designers to solve the usability issues before the summative evaluation and before of the place into market of the device. During the summative evaluation, the safety of the device was confirmed, and additional information for further improvements are collected, both in terms of improvements of the user interface and in terms of improvement for the training provided to users.

The study allowed the designers to observe the learning curve of the end-users and to collect information regarding the safety of the device as perceived by the users and their impression regarding the user experience. In particular, even if the observation of the task analysis led the moderators to think that the learning curve of the users is steep and that there was a sensible improvement of the task performance during the device use, the users reported that the learnability of the device is the aspect that needs major improvements. On the other side, we recall that the users found the learnability of the device still good enough.

The study presented some limits. The first is the numerosity of the participants. While 16 participants are considered satisfactory for the regulatory purposes and are considered sufficient for the determination of the usability issues of a medical device, a greater number of participants could define better the usability aspects evaluated with the UEQ questionnaire. Furthermore, the setting of the test is not representative of the device's real use. The

simulation of the operating room could not provide the simulation complete of the device use environment but represented only the layout of a real use setting. The other environmental conditions like noise, patient presence, and the timing could not be reproduced. Also, the intensive use that is completed during the test is not representative of the real use condition. Even the training is not representative, because the manufacturer intends to provide training before the first use in a similar way to the one completed before the simulated use, but additionally, intend to assist in the first sessions of medical device use.

For these reasons, this study is considered complete in terms of identification of usability issues and terms of confirmation for the device safety, thanks to the worst use condition, but is not considered complete regarding the device user experience. Additional studies should be completed to evaluate user perception during actual use.

REFERENCES

Center for Devices and Radiological Health. (2016). *Applying Human Factors and Usability Engineering to Medical Devices*. U.S. Food and Drug Administration. http://www.fda.gov/regulatory-information/search-fda-guidance-documents/applying-human-factors-and-usability-engineering-medical-devices

Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance.), Pub. L. No. 32017R0745, 117 OJ L (2017). http://data.europa.eu/eli/reg/2017/745/oj/eng

Hopper, A. N., Jamison, M. H., & Lewis, W. G. (2007). Learning curves in surgical practice. *Postgraduate Medical Journal*, *83*(986), 777–779. https://doi.org/10.1136/pgmj.2007.057190

International Electrotechnical Commission. (2015). *IEC 62366-1:2015, Medical devices—Part 1: Application of usability engineering to medical devices* (1st ed.). IEC. http://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/31/63179.html

International Electrotechnical Commission. (2016). *IEC TR 62366-2:2016, Medical devices—Part 2: Guidance on the application of usability engineering to medical devices* (1.0). IEC. https://webstore.iec.ch/publication/24664

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *HCI and Usability for Education and Work* (pp. 63–76). Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Ravizza, A., Lantada, A. D., Sánchez, L. I. B., Sternini, F., & Bignardi, C. (2019). Techniques for Usability Risk Assessment during Medical Device Design. *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies*, 207–214. https://doi.org/10.5220/0007483102070214

User Experience Questionnaire (UEQ). (n.d.). Retrieved September 11, 2020, from https://www.ueq-online.org/

Zhang, J., Johnson, T. R., Patel, V. L., Paige, D. L., & Kubose, T. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics*, *36*(1), 23–30. https://doi.org/10.1016/S1532-0464(03)00060-1.