Are Mobile Devices Ready for Telementoring? A Protocol Design for Randomized Controlled Trials

Andrius Budrionis, Gunnar Hartvigsen, Johan Gustav Bellika
Norwegian Centre for Integrated Care and Telemedicine
University Hospital of Northern Norway
Tromsø, Norway
Email: {Andrius.Budrionis; Gunnar.Hartvigsen; Johan.Gustav.Bellika} @telemed.no

Abstract—**This paper presents a study addressing the usability issues of relatively small touchscreen devices used as endpoints for telementoring. The trial is motivated by the need for systematic knowledge and user experiences on the use of mobile devices for remote supervision of surgeons. Having a stationary computer equipped with relatively large screen and using mouse as an input device in mind, we challenge mobile touchscreen gadgets in order to find out if the same (or sufficient) qualities of mentoring from the mentor's perspective are maintained. The presented study protocol exploits crossover randomized controlled trail design addressing the usability of mobile touchscreen devices for controlling a moving scene (video) and freehand sketching.**

Keywords-telementoring; mobile devices; usability; telestration; platform; RCT.

I. INTRODUCTION

Supervision of medical personnel over distance in order to improve patient outcome has been discussed from 1960s [1]. Technologies changed while the years passed by finally bringing the computational power, required for establishing telementoring session, to the devices surrounding us every day. One could call it the time when the dreams come true - the high availability of medical experts without any dedicated hardware became feasible. The employment of the ubiquitous technology to serve as a mediator between two remote parties pushed the domain to a new dimension. However, is there any proof that the new dimension ensures the same qualities of mentoring? Literature search identifies no interest in investigating the aspects of technology shift in clinical settings. It supports the claim that mobile touchscreen devices were considered and accepted as inevitable technological progress. But, is it the way to go? There are no doubts that mobile gadgets used in random settings are more likely to fail than stationaries, located in surgeon's office. Varying network coverage, battery constrains, interruptions by surrounding people and many other reasons create a fertile environment for unforeseen adverse events. However, not enough credit for the mentioned facts was given in publications so far.

Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) defines telementoring (or teleproctoring) as "a real-time and live interactive monitoring (evaluation) of technique(s) or procedure(s) of an applicant seeking privileges, or a surgeon seeking to certify or document his competence in a specific technique or procedure(s)" [2]. It is a process consisting of two interacting parties – a surgeon (mentee), performing an operation in the operating room and seeking for an advice from the domain expert (mentor) who is not available on site. The paper focuses on the infrastructure to facilitate the interaction between mentor and mentee.

The recent shift from stationary platforms to mobile touchscreen devices has not left telementoring systems behind. While bringing new features and possibilities, it brought new challenges as well. This paper summarizes the protocol for a Randomized Controlled Trial (RCT) aiming to compare different devices and find out whether they can be used as mentoring endpoints on mentor side. In the scope of the paper, platform refers to different hardware employed in a study, not looking into the differences of software platforms.

A telementoring system, developed in Norwegian Centre for Integrated Care and Telemedicine (NST) and deployed at University Hospital of North Norway (UNN), is used in the experiment. Telestration (drawing of freehand sketches over live video) is used in a combination with Video Conferencing (VC) system [3]. Notwithstanding the advantages of live video annotating in actual telementoring session, using the feature in the experiment allows us to compare mouse and touchscreen inputs for mentoring, extending the comparison of the platforms [4].

The paper is structured as follows: after a brief background on telementoring, the gap in research regarding the usability of touchscreen devices for telementoring is stated. Method section covers the study design and scenario, followed by the expected results. The paper is concluded by discussing the advantages of selecting the crossover study design and admitting potential biases and weaknesses.

II. RELATED WORK

Search for research comparing different input devices reported a low interest in the analyzed topic. Baldus and Patterson summarized the reported attempts to measure the differences in performing pointing and dragging actions in still scenes and office environment. Moreover, a comparison of mouse, touchscreen and touchpad was presented while controlling a still scene in a moving environment (vehicle) [5]. A gap in research dealing with usability of different

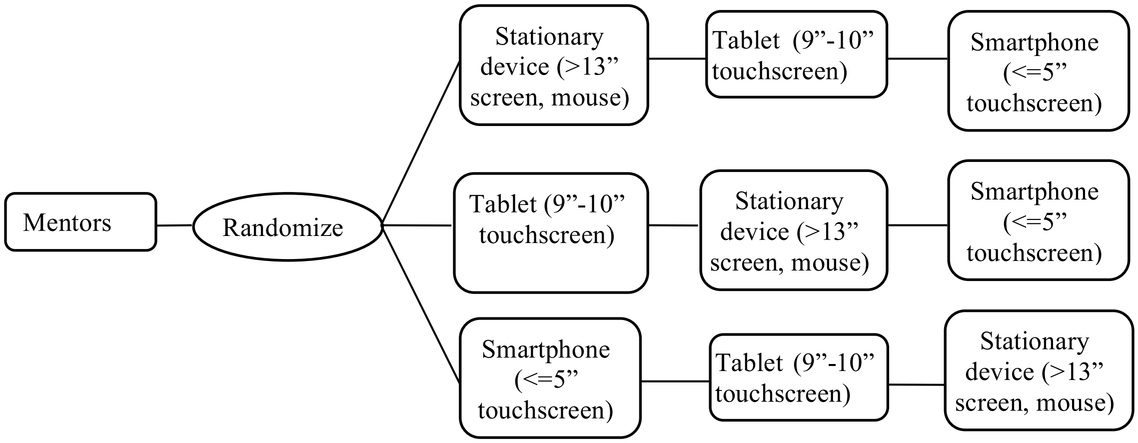

Figure 1. Schematic Trial Design

inputs while controlling a moving scene (video) and freehand hand sketching was identified.

III. METHOD

To produce sound proof on the analyzed topic, a RCT was constructed, employing the crossover study design (Figure 1). Instead of employing the common approach for the RCT (one type of intervention per arm), we let the participants experience all 3 devices in a randomized order. The absence of dependency on the interventions (devices) allows minimizing the number of participants as well as enables the reflection of preferences on the platforms [6]. Due to the high (and increasing) number of mentoring devices, generalization initiative was imposed. The pool of endpoints was divided into three groups based on the screen size, forming the arms of the study:

1. Screen size >13” laptop/desktop computer located in the office of the surgeon, representing stationary platforms. Mouse is used for annotating;
2. Touchscreen size 9”-10” Tablet computer, representing middle-sized mobile devices;
3. Touchscreen size <=5” Smartphone, representing small-sized mobile devices.

Even though the technologies allow using a wide selection of devices on each arm (which would be a typical scenario in a real case), the choice was limited in order to produce consistent and comparable results independent on different hardware on the same arm of RCT. The following devices were selected to represent the platforms:

1. Stationary device – Lenovo X220, I7, Windows 7 equipped with external monitor;
2. Tablet – Asus MeMO Pad, Full HD, Android 4.2;
3. Smartphone – Samsung Galaxy S4, Android 4.3.

Public wireless network infrastructure at the hospital is used for the experiment. All devices run the latest version of Google Chrome web browser as client software to run the telementoring service.

Surgeons at UNN are recruited to participate in the study. As the study employs an imitation of a mentoring session, the inclusion criteria for the participants are not emphasized. The properties of mentoring, observed on the mentee side are not taken into consideration in this study.

Every participant is asked to perform the same mentoring task on all three platforms in a randomized order. After each device, they fill in the questionaire, reflecting their experiences on the mentoring endpoint. Minimum washout period between testing different device is set to 3 days. Results are accumulated on a server side database for further analysis.

In the scope of this particular study, we collect the following data:

1. Mentor response time - duration between the initiation of mentoring session and mentor being present online (Figure 2);
2. Mentor’s interaction with the device – coordinates of annotations, use of pause, resume and zoom functions are logged;
3. Final outcome of mentoring – video and overlaid annotations are recorded;
4. User experiences on every device are recorded by filling in an online questionnaire after each round.

The task for this particular experiment was defined as an imitated surgical mentoring session. The following scenario is being pursued: participants of the experiment are given the devices they will use for mentoring during a regular work day at the hospital. At random order and time, they are notified (by email and text messaging) to connect to the

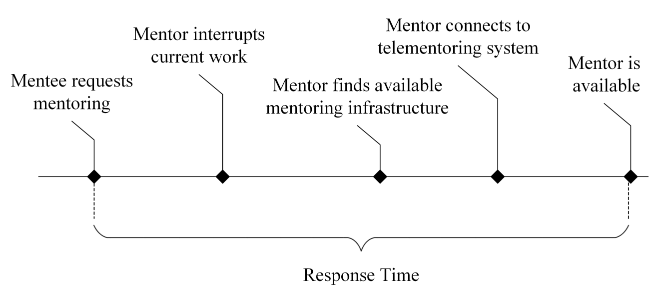

Figure 2. Mentor response time

mentoring service and perform the task. After the mentor is connected, a short video, recorded during laparoscopic procedure, is broadcasted to the device. Participants are asked to identify and mark certain locations in the video (Figure 3).

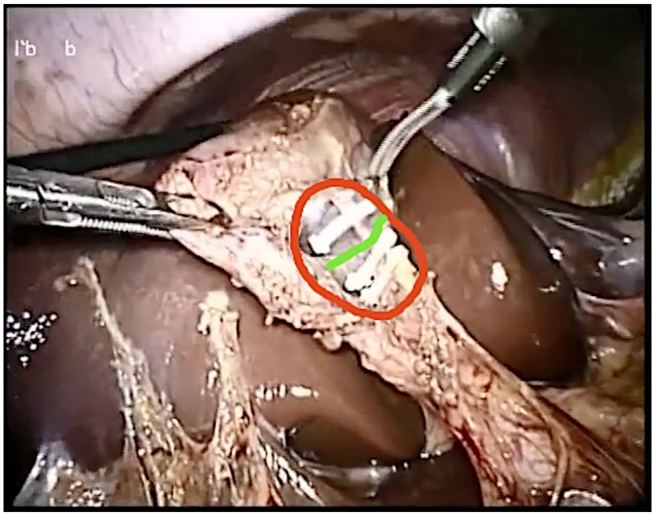

Figure 3. Telementoring task

A schematic view of the trial is depicted in Figure 4. Mentor side of the link facilitates all the functionality required to complete the mentoring task, while mentee part acts as an infrastructure for capturing the progress of the mentor for further analysis.

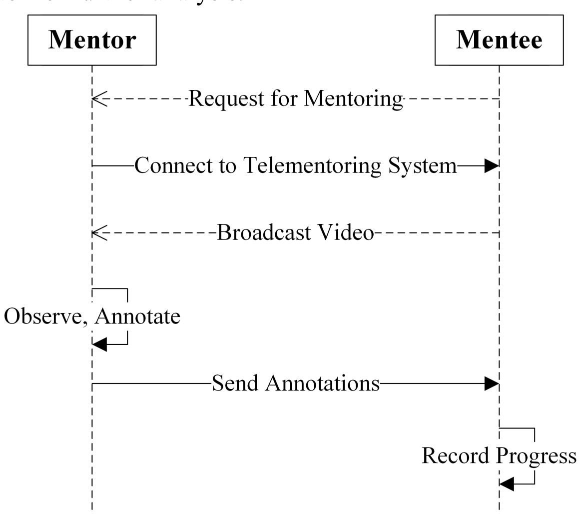

Figure 4. Mentor-Mentee interaction

IV. EXPECTED RESULTS

The paper presents a protocol for the oncoming study focusing on the usability of newly introduced devices. It is addressing the topic, which until now was considered to be natural due to the technological progress. While the move from stationary to mobile platforms is natural in many settings, medical domain deserves a more detailed outlook. The critical scenarios in the domain require in depth research before adopting the new devices.

The study, firstly, looks into the response time of the surgeon on call. It is defined as duration between mentee initiates request for supervision and mentor is connected to the system and is ready to assist (Figure 2). Due to the use of ubiquitous technologies, it should considerably decline, shortening the duration of the procedure (no need to get to predefined mentoring "station", mentor's office to supervise).

Secondly, we aim at studying whether the representation of surgical videos on small screens ensures the same (or sufficient) perception of the progress transmitted from the operating room. The complexity of projecting and perceiving a high resolution video on a small screen are obvious, however, having the technical advances of screen technology and the experience of surgical personnel in mind, the applicability of different sized endpoints for mentoring needs to be tested and evaluated.

Finally, we look at the way the user interacts with the device. In our case, it is either using mouse input or touchscreen in order to produce freehand annotations over live video stream. The study questions if different inputs can generate the same or comparable result, when it comes to accuracy. Is touchpad as good (or good enough) compared to mouse?

It is difficult to answer the postulated questions based on quantitative measurements. Differences of the devices (screen size, input using touchscreen of mouse) may have influence on mentoring process. However, we aim at answering whether the sufficient quality of mentoring is maintained while roaming among the platforms. Moreover, the trade between increasing availability of the domain experts due to the use of mobile ubiquitous devices and higher quality and accuracy, possibly ensured by stationary platforms, is also worth mentioning. No studies report what qualities of the mentoring process are considered sufficient. The results of the study contribute to defining the minimal set of requirements for surgical telementoring systems [7].

V. CONCLUSION AND FUTURE WORK

The paper presented a method to evaluate the use of mobile devices in health care settings with respect to the established technology (stationary platform – control arm in RCT design). It built the fundamentals for further investigations following the presented template. Imitated surgical telementoring session was selected for the case study. The identified gap of knowledge in literature regarding the use of different input devices in clinical settings encourages making the study protocol more generic for applying it in a wide range of settings. Generalization and reuse of the presented approach is straightforward. The main challenge is classification of high number of technological instances to representative clusters.

The selected crossover study design gives a comprehensive comparison of the different platforms. The main advantage of taking this approach is the fact that study could be performed including relatively low number of participants. As the order of using different devices is randomized, every mentor gets a chance to try every device. Eventually, the comparison of results can be performed at the level of individual (using results from testing stationary device as control), minimizing the bias caused by previous personal experience with the technologies or specific technical skills. It supplements the generalized comparison of the results among the different RCT arms [6].

The weak point of the study, possibly having some implications on final results, is the memory effect of the participants. They are asked to perform the same task three times, which will stimulate learning and may introduce bias. To minimize it, the order of the devices is randomized for every participant. In addition, washout periods after each step of the trial are imposed.

We also admit that the results of the study may be case and mentor experience dependent, especially when it comes to perceiving important internal body structures, represented on a small screen. However, this study emphasizes the use of the different devices and experiences of the users rather than mentoring process itself or how it is perceived on the mentee side. Therefore, the mentoring task performed on each device is kept simple to minimize the advantage of more experienced mentors. By including the mentors based on their experience it could improve the outcome; however, an objective measure of surgical skills is complicated due to the number of approaches [8]. In addition, differentiation based on experience would complicate the inclusion criteria for the participants, increasing the numbers of the surgeons to be recruited, as well as producing partitioned results.

Future work, first of all, concerns the analysis of the results from the current study. A series of trials, following the same study design, are planned in order to test and compare the properties of different platforms and video processing techniques used for remote mentoring. The results will form solid fundamentals for the development of telementoring systems.

ACKNOWLEDGEMENT

This research was funded by a Helse Nord grant (ID 5614/HST1025-11) to the Norwegian Centre of Integrated Care and Telemedicine (NST), University Hospital of North Norway (UNN). The study would not have been possible without the support of the Mobile Medical Mentor (M3) project team at NST and collaborators at UNN.

REFERENCES

[1] K. M. Augestad and R. O. Lindsetmo, “Overcoming Distance: Video-Conferencing as a Clinical and Educational Tool Among Surgeons,” *World J. Surg.*, vol. 33, no. 7, pp. 1356–1365, Apr. 2009.

[2] Society of American Gastrointestinal and Endoscopic Surgeons, “Guidelines for the Surgical Practice of Telemedicine,” [Online]. Available from: http://www.sages.org/publications/guidelines/guidelines-for-the-surgical-practice-of-telemedicine/ 2014.01.15.

[3] A. Budrionis, K. M. Augestad, and J. G. Bellika, “Telestration in Mobile Telementoring,” in *eTelemed 2013, The Fifth International Conference on eHealth, Telemedicine, and Social Medicine*, 2013, pp. 307–309.

[4] A. Budrionis, K. M. Augestad, H. R. Patel, and J. G. Bellika, “An Evaluation Framework for Defining the Contributions of Telestration in Surgical Telementoring,” *Interact. J. Med. Res.*, vol. 2, no. 2, p. e14, Jul. 2013.

[5] T. Baldus and P. Patterson, “Usability of pointing devices for office applications in a moving off-road environment,” *Appl. Ergon.*, vol. 39, no. 6, pp. 671–677, Nov. 2008.

[6] B. Jones, *Design and analysis of cross-over trials*, 2nd ed. Boca Raton, Fla: Chapman & Hall/CRC, 2003.

[7] A. Budrionis, K. Augestad, H. Patel, and J. Bellika, “Towards Requirements for Telementoring Software,” *Scand. Conf. Health Inform. 2012*, 2012.

[8] I. Ghaderi, F. Manji, Y. S. Park, D. Juul, M. Ott, I. Harris, and T. M. Farrell, “Technical Skills Assessment Toolbox: A Review Using the Unitary Framework of Validity,” *Ann. Surg.*, Jan. 2014.