

How Trainees Use the Information from Telepointers in Remote Instruction

AZIN SEMSAR, University of Maryland, Baltimore County, USA

HANNAH MCGOWAN, University of Maryland, Baltimore County, USA

YUANYUAN FENG, University of Maryland, Baltimore County, USA

HAMID ZAHIRI, Anne Arundel Medical Center, USA

ADRIAN PARK, Anne Arundel Medical Center, USA

ANDREA KLEINSMITH, University of Maryland, Baltimore County, USA

HELENA MENTIS, University of Maryland, Baltimore County, USA

Researchers have shown both performance drawbacks and benefits of using telepointers or similar display overlay-technologies in remote instruction; however, there is not a clear understanding of why there are these performance effects. This poses a challenge in knowing how and when to successfully use or design telepointing technologies in remote instruction. A better understanding is needed with the rise of remote workers in a wide array of industries from oil rig repair to surgery, and the proliferation of heads-up displays or telecommunications devices to support these future work practices. In this study, we explore how the information conveyed through a telepointer is taken up and acted upon by surgical trainees in a laparoscopic surgical telementoring setting. We collected audio and video data of 12 surgical trainees who performed standard laparoscopic surgical tasks on a physical model under the guidance of a surgical trainer. We investigated both action and talk to determine how the telepointer-based information was used. Our findings reveal three main challenges in using the instructional information conveyed through the telepointer including the trainees' tendency of attending to the telepointer instruction as the primary source of information. We argue that the found challenges are *socio-technical* in nature and require a redesign of the mentoring context as well as the technological tools.

CCS Concepts: **Human-centered computing** → **Collaborative and social computing** → **Empirical studies in collaborative and social computing**

KEYWORDS

Telepointer; socio-technical systems; telementoring; instructional collaborative tasks; information conveyance; shared view.

ACM Reference format:

Azin Semsar, Foad Hannah McGowan, Yuanyuan Feng, Hamid Zahiri, Adrian Park, Andrea Kleinsmith, Helena Mentis. 2019. How Trainees Use the Information from Telepointers in Remote Instruction. In *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3, CSCW, Article 93 (November 2019), 20 pages, <https://doi.org/10.1145/3359195>

1 INTRODUCTION

There has been much work in human-computer interaction and computer supported cooperative work to create tools that support effective distributed work practices [7,11,18,50]. One example of such tools are telepointer technologies - remote pointing and annotation tools.

This work was supported by the National Science Foundation Grants IIS #1422671 and #1552837.

Author's addresses: A. Semsar, H. McGowan, Y. Feng, A. Kleinsmith, and H. Mentis, University of Maryland, Baltimore County, Baltimore, MD, 21250, USA; H. Zahiri and A. Park, Anne Arundel Medical Center, Annapolis, MD 21401, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

Copyright © ACM 2019 2573-0142/2019/November – ART93 \$15.00 <https://doi.org/10.1145/3359195>

PACM on Human-Computer Interaction, Vol. 3, No. CSCW, Article 93, Publication date: November 2019.

Telepointers have been shown to support collaborative interactions by simulating the missing gestural information in a distributed setting [27] as gestural information can clarify or enhance the verbally communicated message in performing a task [15].

Researchers have shown both drawbacks and benefits of using telepointers or similar display overlay-technologies [49,52]. For example, when performing spatial tasks, pointing technologies can impair task performance, while reducing mental workload in a mentoring situation [52]. The performance reduction is assumed to be due to trainee's inability to perform the task on their own because of over-reliance on the instruction provided by the pointer technology. The other assumed challenge of telepointing technologies is the less active exploration from trainees because of overly specific commands from trainers [52]. The fact that telepointer tools are mostly deployed as a single user input gesture tool [11] where typically only the remote trainer can point is another source of problems. It is argued that such tools cause the trainees to passively accept the information provided to them instead of actively acquiring the knowledge [11].

While these assumptions are a start, in order to increase the benefits of using a telepointer and to address possible challenges on the collaboration process, a detailed exploration of trainees' behaviors in taking up and acting upon the given information is essential. Our immediate question is not about the positive and negative impacts of the telepointers, but rather how the information provided by a telepointer is used by the information recipient.

In the following study, we collected audio and video data of surgical trainees performing a standard surgical task while being instructed by a remote trainer using a telepointer. The telepointer enables a trainer to point or draw a freehand sketch over live laparoscopic video. Our approach in using conversation analysis allowed us to explore both the actions and talk around the use of information that is transferred by the telepointer. Based on our analysis, three main challenges are revealed in using and acting upon information provided by the telepointer. These challenges include trainees' tendency of attending to the telepointer instruction as the primary source of information, parsing sequences of information provided by the telepointer, and insufficiency of the information provided by the telepointer in proceeding with manipulations having significant consequences. We further argue that these challenges are neither merely technical nor solely social, but we face a socio-technical problem. Finally, we present socio-technical implications for each identified challenge.

2 BACKGROUND

2.1 Tele-pointing

Telepointing technologies have a long history within the CSCW community. In remote group-work settings, they are typically implemented as a tool for the remote person, replacing his/her traditional cursor with a pointer that replicates the person's interactive hand movements [10]. Telepointers use in remote collaboration rose in the early 1990s by introducing systems such as Agora [34], Commune [5], GestureCam [33], MMM [4], and DOVE [45], and it is still a focus of research in the community [1,11,27]. They were created to enhance the effectiveness of distributed collaboration through explicit visual feedback. The systems enabled a remote expert to guide the actions of a local worker through visual instructions. The ability to guide actions through visual instructions allows for explicit instruction, and in fact, pointing on the shared view could represent the remote person's instructive gesture, which would have otherwise been missed in a distributed setting [27]. These technologies have been studied for mentoring purposes, specifically, when a trainee needs to be guided by a remote trainer.

Research results have so far demonstrated the benefits of using telepointers in remote collaborations [19,31]. For instance, it has been shown that the telepointer increases joint action coordination, allowing the group to see when actions begin and end, and enabling people to anticipate one another's actions [19]. It has also been demonstrated that the local worker's ability

to accomplish a novel task is improved when being instructed by a telepointer as compared to a standard video-mediated condition [31]. These results indicate that the telepointers can facilitate the group process and enhance the novice local worker's skill of performing the task.

Along with the benefits, researchers have uncovered some possible challenges with the use of telepointing technologies and in kind have suggestion for further design considerations. With an experiment of robot-assembly task with pairs of students, Fussel et al. [15] showed that the cursor pointing tool did not improve the students' performance over the scene view only condition. They hypothesized that performance did not improve due to the limited functionality of the pointing tool, in that it only supported pointing, but not any representational hand gestures. This would include such crucial instructional information as the orientation and the direction of movement. In a different study, Fussell et al. [14] found that additional gestural information is required to be provided by the telepointer to demonstrate the actions and increase performance. However, Kim et al. [29] found that with the use of a pointing technology, verbal communication was still the main modality for conveying crucial orientation information because the telepointer is unable to represent such information. The importance of a combination of verbal and pointing instructions, also, was a finding of early studies of embodied references in workplace interactions within which verbal and pointing instruction, hand in hand, provide the context to interpret the deictic instructions [24,32]. In still other work, a system that only provides a remote-worker with a pointing tool was found to impair the balance of communication structure, as the local worker passively follows the pointer without contributing his or her own input [11]. All these studies refer to missing functionalities, which represent some technical issues with the telepointing technologies. Compared to the technical design suggestions, social implementation considerations in how collaborators use the telepointers are not much covered in literature. Providing social clues regarding how to better use the telepointer in cases could call for less effort and cost as compared to technical solutions.

Although prior work identified challenges of using telepointing technologies, the challenges were not the focus of the study. Instead, they were discussed as a way to explain why the technology may have negatively affected the group process or outcome. In the following study, by analyzing the detailed talk and actions of an instructor and learner, we are deeply focused on the challenges to uncover both possible technical and social solutions with the current telepointing technologies.

2.2 The Rise of Tele-instruction

With the rise of high-speed networks (e.g. 5G network) over the past two decades, tele-instruction - providing instruction from a distance - has become an active field both in research and in industry. The primary tele-instruction systems were mainly based on simple audio and video communications [13,28]. Later various interaction and visualization tools were developed and studied to reinforce guidance over video-mediated systems. Examples are remote pointing [44], remote annotations [16], remote hand gesturing [30], and mixed reality tools [25]. The common need of tele-instruction tools even has made smart device producers to build applications supporting remote instructions. Recently, Microsoft company released the mixed reality Dynamic 365 Remote Assist application with the unveil of Microsoft HoloLens version 2 with the aim of helping technicians to achieve guidance from a remote expert [40]. The company recently announced that the same application will be available for Android mobiles with the purpose of providing the same support through every-day common smart devices such as mobile phones [53]. Mobile mixed/augmented reality applications also have been investigated by researchers to support tele-instruction. For instance, Datcu et al. [9] presented a distributed collaborative Augmented Reality system running on the smart-phone of a local worker showing that with the use of a less intrusive device (i.e. mobile device), the collaboration was enhanced while avoiding the limitations of more obtrusive tools (e.g. Head Mounted Displays).

With the advances of tele-communication technologies, more recently, tele-instruction applications and tools have been applied to various applications such as telemedicine [27], telesurgery [8], and remote technical tasks and repairs [41] for supporting remote instructions. In telemedicine, tele-instruction applications are used in various cases including for distributed medical experts to consult on patients' captured radiology images and videos [35], and for emergency medicine remote support [43]. In technical tasks, a remote technician uses tele-instruction tools to guide a local technician in performing tasks such as assembly, troubleshooting, and repair.

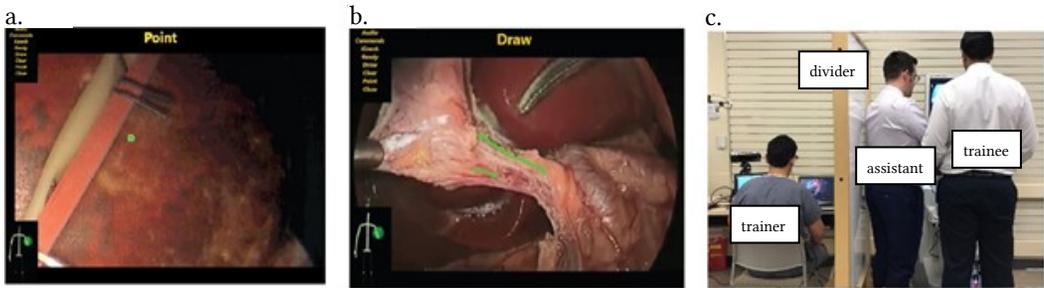


Fig. 1. The telepointer: (a) telepointer user interface in pointing mode - the pointer is the green dot on the anatomy; (b) drawing mode - green lines trailing green dot; (c) Trainer on left using the telepointer and speaking over the divider to the trainee.

As tele-instruction applications expand to various domains and telepointers are potentially common tools for improving the quality of tele-instruction process we need better understanding of the challenges in using them in order to devise better designs as well as identify best practices for use and deployment. In the following study we investigated these challenges in the domain of surgical telementoring where an experienced surgeon remotely shares knowledge to a less experienced local worker in performing minimally invasive laparoscopic surgery tasks. In laparoscopic surgery, as opposed to open surgery, a long fiber optic camera provides the surgeon with a view of the operative field on a display. This method of surgery is growing in popularity as it poses a lower risk for infection and decreased time to recovery for the patient. However, a significant challenge in learning how to perform a surgery laparoscopically is how to 'see' the anatomy in the display as well as how to manipulate the specialized tools in the confined space [37]. In order to expand the use of laparoscopic surgical tasks, the surgical telementoring has been promoted. For instance, the Society of American Gastrointestinal and Endoscopic Surgeons (SAGES) has introduced telementoring as a mechanism of improving laparoscopic training and increased its access for patients [48]. Through this mechanism, an expert laparoscopic surgeon is enabled to remotely guide another experienced surgeon on new tools and techniques. This promotes global exchange of knowledge with the purpose of improving patient care in remote regions. In current telementoring practices, the laparoscopic video along with the view of the local surgeons' actions outside the body is lively transferred to remote experts to observe and advise accordingly [46,47].

3 METHOD

The study was performed in the Department of General Surgery in the Anne Arundel Medical Center, Annapolis, MD. We obtained Institutional Review Board Approval from the University of Maryland, Baltimore County. Informed consent was obtained from all participants before their participation.

3.1 Telementoring System and Setup

A system was developed in order to study the conveyance of expert surgical knowledge in collocated or distributed interactions. It enabled a trainer to point or draw using free-hand sketches over laparoscopic video for trainees to see (Fig.1 a & b). The pointing and drawing functionalities provided in our home-built system are the common main features incorporated into the majority of telecommunication technologies [29,31,45], and in current telecommunication systems used in surgical telementoring such as the VISITOR1® (KARL STORZ, Germany) and Skype (Microsoft, USA), or telemedicine such as VSee (VSee, Sunnyvale, CA).

In order to simulate the use of the system in a sterile operating room, the Microsoft Kinect sensor version 2 (Microsoft Corporation, USA) was used as a mechanism of touchless interaction [42]. Developed in C# for the Windows Operating System, the application is a transparent window that can be overlaid on any screen or other application. Implementing Microsoft.Speech API and body tracking, it uses either hand movements or audio keywords, or a combination of both to call upon the different functionalities [38], such as a pointer for referencing or a freehand drawing tool (Fig. 1). The Kinect detects the movement of the trainer and listening for verbal keywords. In pointing mode, the user moves the hand to control a small green circle, which acts as a pointer. In drawing mode, the user closes their hand to draw and then stops drawing by opening the hand. When the pointer is not needed, the user can either put down their hand or give the voice command “Kinect end” to turn off the pointer.

To simulate the distributed condition of a telementoring environment, a divider was used between the trainer and the trainee, so that they could not view each other (Fig. 1.c.) but still hear one another’s voice. An assistant held and maneuvered the laparoscope for the trainee while performing the task. The gesture/voice interaction device was used in both conditions in order to control for changes in use due to the mechanism of interaction.

3.2 Study Design

In a prior session, the trainer was introduced to the telepointer system. He was informed with the telepointer functionalities including pointing, drawing, activating, and deactivating the system. He was also trained how to enable each functionality.

Before the study, the trainees and trainers completed a demographics questionnaire, which included information on their surgical experience. The trainees were asked to work on a 6 to 10 minute simulated laparoscopic task under trainer guidance. The Park Trainer (Stryker Corporation, USA) was used for the simulated laparoscopic tasks. The Park Trainer consists of a housing unit for physical anatomical models, a flexible shield with openings for the laparoscopic camera and instruments to be inserted, a standard laparoscopic camera with light source using the Stryker computer system, and a standard laparoscopic monitor on an adjustable arm at the top.

The task was selected based on a hierarchical task analysis of the laparoscopic cholecystectomy procedure (removal of the gallbladder). The task required both skills of anatomical structure identification and instrument manipulation. The task was performed on a validated laparoscopic training physical model [38], including four subtasks as follows: (1) mobilizing the cystic duct and the cystic artery which is dissecting the peritoneum to separate the artery and the duct using marylands or scissors as instruments, (2) clipping the cystic duct using clip applicators, two clips on top and a clip on bottom as shown in Fig. 2.b., (3) clipping the cystic artery using clip applicators, two clips on the top and a clip on the bottom as shown in Fig. 2.b , and (4) cutting the cystic artery and the cystic duct in between the clips applied (Fig. 2).

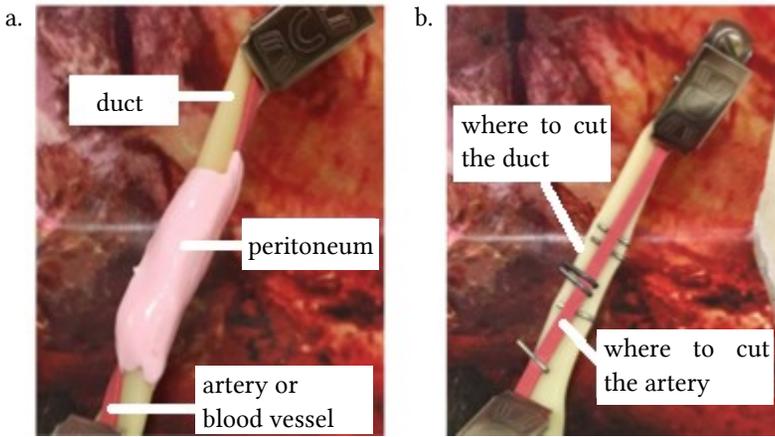


Fig. 2. Task models used in the study: (a) Used for subtask 1 - mobilizing the cystic duct and artery; (b - without the staples across the structures) Used for subtask 2 - clipping the cystic duct; and subtask 3 - clipping the cystic artery; (b - as shown) Used for subtask 4 - cutting the cystic artery and duct.

Table 1. Trainees' Experience

Trainee	Post Graduate Year	Laparoscopic Experience	Gender
1	PGY-2	minimal	Male
2	PGY-2	minimal	Male
3	PGY-3	intermediate	Male
4	PGY-3	intermediate	Male
5	PGY-2	minimal	Male
6	PGY-1	minimal	Male
7	PGY-1	minimal	Female
8	PGY-1	minimal	Male
9	Research Fellow	intermediate	Male
10	Intern	minimal	Female
11	surgical fellow	advance	Male
12	PGY-3	intermediate	Male

3.3 Study Participants

The participants (Table 1) were recruited from the Department of General Surgery in the Anne Arundel Medical Center, Annapolis, MD. A total of 12 surgical trainees (10 male), including 1 surgical fellow, 1 research fellow, 1 intern, and 9 surgical residents (3 PGY-1, 3 PGY-2 and 3 PGY-3 - PGY refers to Post Graduate Year in medical school) were recruited. All the participants had at least some minimal experience in performing laparoscopic tasks. One attending surgeon was recruited as the trainer.

3.4 Data Collection

Video recordings enable more detailed, reflective and systematic analysis of the unfolding actions of the participants than in-the moment observations. Since we were interested in understanding the trainees' use of the information in very detailed, instantaneous and in-situ manner, we collected data including audio and video recordings of trainees and trainers while performing the task.

An external camera was set up to the left of the trainer to capture the trainer's hand movement and the view of the telepointer on the trainer's screen. Another camera was setup to the right of the trainee to capture the trainees' movement and the laparoscopic view. The laparoscopic view was screen recorded on the Laptop as well.

3.5 Data Analysis

As a sum, 92 minutes of videos were recorded in the data collection phase. The entire recordings were used for the data analysis. After collection, a combined video view was created that contained (1) the external camera video of the trainer's movement and his screen (over the shoulder), (2) the external camera video of the trainee's movement and his or her laparoscopic screen (over the shoulder), and (3) the screen recording of the operative field (via the laparoscopic camera). The picture in picture format of all three sources of video recordings juxtaposed next to one another can facilitate the analysis of the recurring actions and talks concurrently.

Our focus for the analysis was around uptake and use of the information provided to the trainee through the telepointer as an instructional collaborative tool. A detailed transcription containing both trainers' and trainees' action and conversation was made of the picture-in-picture videos from the entire corpus by the second author. Then the analytical practice utilized by Heath et al. [20] and later used by Mentis et al. [39], ethnomethodologically-informed Conversation Analysis [22], was conducted by the first author. This practice includes a detailed reflective analysis that captures utterances, pauses, overlaps, intonations, and visible actions. Not all transcriptions were chosen to be parsed in the Conversation Analysis format. Instead, the first author watched the whole recording, second by second, and then extracted the moments relevant to the uptake of information provided by the telepointer. These moments were written in vignettes. We identified recurring patterns of behavior that build the main themes of instruction uptake and use. The most representative examples are provided in the results section.

The authors have sufficient background understanding of the professional task. The last author of this paper has been working in this particular area of surgery for 7 years; under her guidance the first author has been studying this surgical task for two years and the third author has been conducting studies around these particular surgical tasks for 4 years. Two additional authors on the paper are laparoscopic surgeons.

4 RESULTS

We identified three main challenges in the ways that the trainees take up, use, and act based on the conveyed information. The first challenge refers to the extent the trainees prioritize the telepointer instruction over the verbal form of instruction. The second challenge relates to how trainees treat a sequences of telepointer instruction. The third challenge describes how trainees use the telepointer instructions in performing manipulations with significant consequences.

4.1 Attending to the Telepointer Instruction as the Primary Source of Information

In a complex critical training environment such as laparoscopic surgical training, trainees are required to manage multiple cognitive and physical actions while they are exposed to multiple sources of information at the same time. In such an environment, the trainee needs to integrate different skills, knowledge and behaviors concurrently. A telepointer is a new source of information added to all the information they were already exposed to. In this case, the trainees show a tendency to attend to only a single source of information resulting in missing the information conveyed in another medium. The trainees show a tendency to prioritize the sources of information resulting in an ordered list of information sources. By applying this approach, they either use the telepointer as the primary source of information and the verbal instruction as the secondary or vice versa. When a trainee is overloaded with the given information, they may fail to share their attention with different sources of information concluding in paying all of their attention to their primary source of information to the detriment of the others. Such tendency is revealed mostly in two cases.

The first case is the moments that the instruction provided by the telepointer is wrong or it is different from the verbal instruction while both conveying the same type of information (e.g. location of an action). For instance, one major type of instruction provided by the use of the telepointer is guiding the trainee to direct the instrument to the correct anatomical target location. This instruction helps the trainee to locate the correct point of action in using the instrument by providing visual indicators on the exact desired point of application. In multiple sessions, some instances of pointing on the wrong spot occurred. For example, the trainer mistakenly points on a tissue, although he verbally instructed the trainee to direct the instrument to a different tissue. This was either due to the system's challenges or the trainer not being as exact as necessary. In the following example, we show that the trainee (P8) acts upon the wrong instruction provided by the telepointer.

In this vignette, we see the trainer mistakenly ended up pointing to the duct instead of the artery (in line 3). The reason might be that he focused on giving the verbal instruction and he does not fully attend to where he is pointing. The trainee (P8) seems to prioritize the telepointer instruction over the verbal instruction (in line 4). The reasoning behind this behavior could be the trainee's incapability in managing all the information he is exposed to as he has less experience in handling the cognitive load of this surgical task. In such a situation, he reacts to the telepointer instruction because that is the most salient piece of information in front of him. And so, the telepointer instruction as a visual source of information supersedes the verbal instruction in catching the trainees' attention.

The second case that reveals the use of telepointer as the prominent source of information is the moments that the telepointer and verbal instructions refer to two different types of information (e.g. telepointer shows the location and the verbal instruction is about the orientation of the instrument).

```

1   trainer  then you are going to come with your dominant hand and
           you're gonna start creating a window a little divot right
           there
           ((points at the location))
2   trainee  ((moves the instrument towards where the pointer points at
           while tips are turned outwards of the tissue))
3   trainer  turn your tips inwards so you are pointing to that divot
           with tip of your Instrument,
4   trainee  ((still approaches the tissue with tips outwards))
5   trainer  nope nope no no no
6   trainee  ((pulls back the instrument))
7   trainer  <turn the tips of your instrument inward>
8   trainee  [((turns the tips inward) ]
9   trainer  [keep turning (0.1) yup so the tip is pointing to that
           divot]
           start spreading in there (0.7) gently.
10  trainee  ((starts dissecting where the pointer points at))
11  trainer  ((moves the pointer to the right side of the peritoneum))
           [right there]
           [((fixes the pointer at a point on the right side of the
           peritoneum))]
           To u::hh open up [this little space ]
           [moves the pointer up and down, back and
           forth on an invisible line between two close points
           crossing the spot he was pointing at]
12  trainee  ((moves the instrument towards where the pointer pointed
           at while the tips were towards outside of the peritoneum))
13  trainer  turn your tips please,
14  trainee  ((rotates the instrument so that the tips turns towards
           the peritoneum))
15  trainer  yup keep dissecting that tunnel (0.5) and work your way
           up.

```

Here, P6 starts the task by being instructed to dissect the peritoneum. The trainer uses both verbal instruction to guide him to use the right orientation of the tip of the instrument (in line 3), and telepointer instruction, to show him the right location of dissection (in line 1). What we see is the trainee ignoring the verbal instruction regarding turning the tips and following the telepointer location (in line 4). Thus, we notice that the telepointer is the primary source of information for the trainee at this moment. In response, the trainer prevents him from moving further while holding the instrument in the wrong orientation by giving an explicitly emphasized hindering verbal command, “Nope no no” (in line 5). This moment is when the trainee again turns his attention to the verbal instruction.

This behavior could be due to the trainer and the trainee’s different approaches in using the telepointer instruction. From the trainer’s point of view, the telepointer might be an aiding technology to facilitate the verbal instruction. Meaning while the telepointer does not provide prominent information at some moments, it is still activated. However, the trainee might use the

telepointer as the main source of information. Therefore, this mismatch in perceiving the role of the telepointer must be addressed by an integrated design and training solution.

4.2 Parsing Sequences of Information Provided by the Telepointer

Trainees need to remember a lot of information when they are instructed including the type of manipulation, the location of the manipulation, the type of instrument, the orientation of the instrument, etc. Furthermore, they need to integrate what they see, what they hear, and their background knowledge. With this, being provided with a sequence of instructions might be challenging in terms of the amount of information needing to be remembered.

One common situation that arose is when the trainer wants the trainee to simultaneously perform different manipulations by each hand on two different spots. Two approaches may be applied to provide the instruction that includes two simultaneous actions. The first approach is to give all the information needed for both actions at the beginning.

```

1  trainer  I want you to pickup the tissue< u::hh righ::t
           ((moves the pointer to the right side of the peritoneum))
           right around there somewhere
           >pick that peritoneum< and you're gonna dissect
           with your non-dominant hand you're gonna grab and retract
           ((holds the pointer on the right side of the peritoneum))
           and you're gonna dissect underneath (0.4) here
           ((moves the pointer to an upper spot))
           to find the blood vessel and separate the peritoneum off
           the blood vessel. so go for that.
2  trainee  ((brings down the instrument toward the peritoneum))
3  trainer  if it helps to go the opposite direction do that. you can
           also grab the peritoneum towards you
           [from this side and use your other hand      ]
           [((points at the left side of the peritoneum))]
4  trainee  ((grabs and [pulls the left side of the peritoneum with
                                           left hand]))
5  trainer  [pull (0.1) yeah                    ]
           with your left hand (0.5) pull a little bit
           [and begin to kinda]
6  trainee  [((moves the instrument in the right hand toward the left
           side of the peritoneum underneath where he grabbed by
           non-dominant hand))]
7  trainer  >not there
8  trainee  ((pulls back the instrument in his right hand))
9  trainer  [Actually right arou::nd there you should be dissecting]
           [((moves the pointer towards the right side of the
                                           peritoneum))]
10 trainee  ((moves the instrument toward where the trainer points
           at))
11 trainer  Yup.

```

In this case, the ultimate goal is to find the blood vessel by dissecting the peritoneum covering it. The trainer starts by explaining the goal and one way of accomplishing it in line 1. The approach is to grab the right side of the peritoneum and pull it with the non-dominant hand (i.e. left hand) and dissect underneath using the right hand. In line 3, he offers another approach which is still grabbing by the left hand and dissecting by the right hand but on different locations. However, the pointer only points at the grabbing location (in line 3). While the trainee, P11, begins to grab the correct location using the left hand (i.e. left side of the peritoneum), he moves the instrument in his right hand towards the same location where the

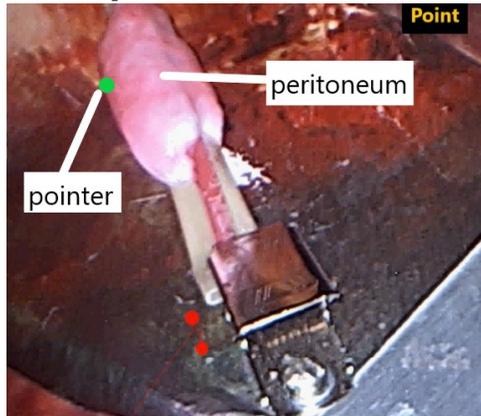
pointer is fixed at (in line 6). That is when the trainer intervenes by saying “not there” (in line 7) and moves the pointer towards the right side of the peritoneum where he wants the trainee to dissect.

In this example, we see that the trainee uses the immediate information provided to him. Although in the beginning, the trainer is pointing at the location to dissect, later, when the trainee starts dissecting, he does not act in response to the visual information showing where to dissect. Instead, the trainee interprets the location to grab as the location to dissect.

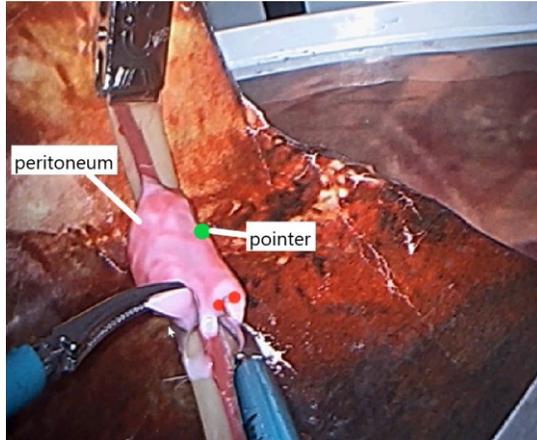
In such situations, we saw that giving the entire sequence of instructions including the two needed manipulations and two locations and then holding the pointer at the location desired for one of the manipulations might mislead the trainee.

An alternative approach is to break down a sequence of instructions into separate instructional moments. With that, the trainer waits for the trainee to perform the first manipulation along with pointing at the desired location, and then moves the pointer to the location of the second manipulation asking to perform the second manipulation. Below is an example of a trainee being instructed by the latter alternative approach.

```
1  trainer  [I want you to use your nondominant hand and grab right
              there]
              [(moves the pointer to the left side of the peritoneum)]
```



```
2  trainee  Where I am pointing (0.5) you see me?
              Yes
              ((moves the instrument in his left hand towards where the
              pointer points at))
              ((grabs the peritoneum on where the pointer points at))
3  trainer  A::nd (1.2) put tractions towards you
4  trainee  ((pulls where he grabbed towards himself))
5  trainer  [And the::n u::hh dissect a little bit just medial]
              [(moves the pointer to the right side of the
              peritoneum)
              ]
              [right there]
              [(fixes the pointer at a point on the right side of the
              peritoneum)]
```



To u::hh open up [this little space]
 [moves the pointer up and down back and forth on an invisible line between two close points crossing the spot he was pointing at]

- 6 trainee ((moves the instrument towards where the pointer pointed at while the tips were towards outside of the peritoneum))
- 7 trainer turn your tips please
- 8 trainee ((rotates the instrument so that the tips turns toward the peritoneum))
- 9 trainer yup keep dissecting that tunnel (0.5) and work your way up
- 10 trainee ((starts dissecting the area between the two points the pointer pointed at))

Here, the trainee is asked to perform the exact same two manipulations; grabbing with the left hand and dissecting with the right hand. The difference in this situation is that the trainer splits the two sets of instructions. First, in line 1 he asks the trainee to grab the peritoneum while pointing at the desired location. By so doing, he provides the two main pieces of information needed for accomplishing the task, the manipulation type and the location, right when the trainee carries out the task. After completing the first manipulation by the trainee, the trainer starts asking for the second manipulation along with pointing at the location with the telepointer in line 5. In this situation the immediate pointing location the trainee sees shows the desired spot of the action he is to maneuver through. Our examples show when a sequence of manipulations are needed to be performed, instructing the trainee without pause between the different manipulations may lead to confusion. The fact that the telepointer is designed to only convey a single focal point on the view at a time is the problematic factor as the trainees are sometimes asked to perform two manipulations on two different locations in a continuous form of instruction (without pause). Indeed, telepointers are considered to convey the following types of information: (1) where to focus one's attention for description (2) where to perform a single action, and (3) how to perform a single action. In our example, the trainer fails to provide the instructions for two different location at a time because the telepointer is not designed to convey that type of information. Thus, in order to use the information that can be gleaned from a telepointer, the method of instruction needed to be changed to that of a sequence of actions. In other words, to break down the simultaneous actions into two separate instructions, which then allowed for the use of the single focus/single performance information that the telepointer can assist.

4.3 Insufficiency of the Information Provided by the Telepointer in Manipulations with Significant Consequences

A surgical task is a combination of several different manipulations. In our examples, a task includes manipulations such as grabbing, dissecting, cutting and clipping. However, some manipulations might carry more significant consequences compared to others. For instance, the consequence of peritoneum grabbing or dissecting might not be as significant as cutting a blood vessel. Such differences lead the trainee to behave differently towards telepointer instructions. Below is an example from P11 being instructed to cut the blood vessel.

```

1  trainer  so, now you're going to cut between the upper part of your
           second
           >[so you're gonna cut right there.]
           [((points at a location on the blood vessel between the
                two clips))]
           take your scissors and cut that.
2  trainee  if it helps to go the opposite direction do that. you can
           also grab the peritoneum towards you
           [from this side and use your other hand          ]
           [((points at the left side of the peritoneum))]
4  trainee  ((grabs the scissors and takes them into the laparoscopic
           view))
           ((puts the scissors tips on the spot the telepointer
           points at))
           right here?
5  trainer  yup.slide in (2.0) go fire.
6  trainee  ((cuts the vessel at the pointed location))

```

In this example, the trainee does not simply follow the location of the pointer for cutting. Instead, he asks for a verbal confirmation in line 4. Once the trainer confirms the correct positioning of the scissors tips in line 5, then the trainee starts cutting. This same behavior (i.e. asking for confirmation before performing the manipulation) is observed for most of the trainees in applying clips on the tissues and in cutting. However, this is not the case for the grabbing and dissecting types of manipulations. An example is lines 9 and 10 in the first vignette of subsection 4.2. The trainee immediately follows the pointer location for dissecting the peritoneum without asking for confirmation.

Unlike our first set of examples in section 4.1 - that the trainees fully attended to the telepointer instruction ignoring the verbal instruction - in this last example, before they make a significant, unalterable action and even when the telepointer points at the right location, the trainee asks for clarification. The trainee avoids acting immediately in response to the telepointer instruction by asking for verbal verification. This observation leads us to the idea that the trainees act differently upon the telepointer instruction when they carry out a manipulation with a significant consequence.

The other interesting behavior in this example is the use of the instrument for pointing at a location by the trainee in line 4. Although such behavior was rarely observed from the trainees, in most cases where they carried out a manipulation with a significant outcome, they asked for confirmation about the exact location of the action by pointing with the surgical instruments themselves. This indicates three facts. First, the need for trainees to be able to visually refer to a location on the shared display. Second, the imbalanced structure of communication between the trainer and the trainee - apart from the moments the trainees ask for clarification with regards to the location of performing a significant action, the trainees do not actively participate in the discussion. Third, the tendency of asking for confirmation in performing a manipulation with

significant consequences might not occur in a simulated less critical experiment/setting. This study explores a real world work setting. In this environment, all tasks are not the same in terms of how critical they are. This should be noted in any real world work setting that incorporates the telepointer as an instructional technology.

5 DISCUSSION

In this study, we investigated the patterns of actions and talk that resulted from the use of a telepointer as an instructional guidance tool in a surgical telementoring setting. We were interested in unfolding the actions elicited upon telepointer-provided instructions, in order to gain a better understanding of the challenges of using current telepointer design. We revealed three main challenges of using a telepointer technology, specifically in the domain of remote mentoring in surgery. More importantly, we have found that the solutions to the challenges are not simply technical (e.g. found in [11,14,15]) or solely social, but could be addressed in an intertwined social and technical manner. Technically, we need to offer better, more effective, tele-pointing technologies. However, only presenting the technology to the users is not enough. In order for collaborators to know how to convey and perceive the information through the telepointer, a train-the-trainer phase on how to use the technology in the proper way is needed [2]. Otherwise, the telepointer instructions are sometimes successful and other times unsuccessful as we observed in our data. Indeed, we did train the trainer how to use the *functionality* of the technology. But, what is missing is training on how to utilize the tools with regards to trainees' way of uptaking the information. This is about what knowledge the trainer needs to understand how communication and training occurs, and how to utilize the tools in order to properly convey that information.

Here, we face a *socio-technical* problem. A socio-technical problem is a “problem with social, organizational and technical processes interacting to produce the phenomenon to be studied” [6,51]. For a socio-technical problem the solution is not merely about better system design, but it is a multi-tiered integrated solution grounded in both social and technical aspects. Because of the challenges that emerge by inserting the technology (i.e. telepointer) into a social context (i.e. mentoring), we state that the observed challenges we have observed are of a socio-technical manner. In order to manage socio-technical challenges, intertwined social and technical implications must be considered. In the following sections, we discuss both socio-technical implications for each challenge we found in using the telepointer.

5.1 Attending to the Telepointer Instruction as the Primary Source of Information

We saw that in situations within which the local workers were overloaded with information, they tended to prioritize the information sources by paying the majority of their attention to one source of information. In this situation, when the local workers attend only to the information provided by the telepointer as the primary source, and this information is accidentally pointing at the wrong location, the trainees understandably fail carrying out the action on the correct location.

The first technical solution that can reduce issues such as accidentally pointing at the wrong location is to ensure there is a quick mechanism for turning on and off the telepointer - in essence only using it at specific times of instruction [3]. However, having no or difficult commands for turning off the telepointer may hinder the remote expert to deactivate the pointer. This seems like a simple and straightforward technical solution. And this technical solution had been suggested by Heath et al. [21] as well. In the pointing system they proposed, the option of turning off the pointer was not provided. This caused the remote person to unintentionally point on locations because they were not able to turn off the pointer when pointing was not needed. In our system, the option of deactivating the pointer was available for the trainer. However, this technical solution solely did not work in our study. Although an easy

command for turning off the telepointer was provided (removing one's hand or saying "Kinect end"), this did not prevent the occurrence of a trainer continuing to hover over a point when the instruction had moved on. This comparison shows powering the systems by incorporating better technologies is not sufficient. In fact, our argument is not that there is a technical solution that we have not thought of. Our argument is that in order to take advantage of that technical solution, there is a required social context in which that technology is used.

Here, the social challenge might be a disconnect between the foci of the dyad which could be a communication problem which is grounded in the social interaction. The challenge is that the remote expert might not attend to where they are pointing because they are focused on giving the verbal instruction. On the other side, the local worker tends to treat the telepointer as the primary mode of communication as opposed to the verbal channel and that prioritization would be overlooked by the remote expert. Informing the remote expert about such tendencies can inform them to work more carefully with the pointing technology. If a remote expert is trained that a mistakenly pointed gesture may trigger a local worker's error in performing the task, they could continuously pay attention to where the pointer is located over the course of the task and more closely observe the trainees' actions. The remote expert should be taught not to point at a location unless that is the exact spot they want the local worker to attend. This also stands for the cases where the verbal information is so important. The remote expert needs to be trained that the pointer should not be used when they provide important verbal information because the local worker tends to ignore the verbal instruction. They should be taught that decoupling verbal instruction from telepointer instruction, or ensuring that the telepointer points at the right location could prevent errors. Thus, in training-the-trainer, it is not to teach the trainer how to use a specific technology, but rather to train-the-trainer in how trainees use the information provided to them through the telepointer.

5.2 Parsing Sequences of Information Provided by the Telepointer

The majority of the developed telepointing technologies [10,14,15,19] overlay a single hand movement on the shared view. However, as we observed, the local worker might be asked to use both hands to perform manipulations on more than one location at a moment. This is where the pointing technology can partially provide the spatial information. This technical gap can justify the need for a pointing technology which is able to represent instructions for both left and right hands. This can be implemented by triggering an input (e.g. mouse button) that changes the pointer shape and color to represent the instruction for the other hand. For such implementation the question is how adding colors or shapes as information carriers might be cognitively overloading for both parties. The other design solution is to project the remote expert's hand gestures on the shared view so that the local worker can observe both hands placement on the work field view [30]. However, such a design may cover much of the field background information that is critical for complex tasks such as surgery. This issue could be alleviated by incorporating translucent hand gesture images [26,36]. Furthermore, the projected hand gestures might not provide precise information for the small spot a manipulation must be carried out on.

However, with incorporating all possible technical improvements there is still social implication that must be taken into account. Providing a sequence of instructions at a time might be challenging for the local workers as they would be required to remember a lot of information. As we indicated in our results section in the second vignette of section 4.2, it was observed that the trainee did not perform any errors when instruction was presented in a step-by-step format. Here, the trainer waited to provide instruction of the second hand until the trainee completed the manipulation instructed to do with first hand. This waiting behavior may be considered when teaching the remote expert how to instruct using the pointing technology. They could be taught to break a sequence of instructions into pieces to reduce errors that may

have been caused by the cognitive overload from presenting too much information at once. By applying this strategy, the remote expert gives all the required spatial information at the moment of action, to decrease the likelihood of erroneous action by the local worker.

5.3 Insufficiency of the Information Provided by the Telepointer in Manipulations with Significant Consequences

In starting an action with significant consequences, local workers might look for more clarification than only the telepointer provided information. This strategy in asking for confirmation regarding the location of actions with significant consequences (e.g. cutting and clipping in surgery tasks) is not only encouraged, but also should be facilitated through technology. In order to assure the right location for the critical manipulations, we observed that the local trainees used their instruments (e.g. scissors, and clip applicators) to refer to a location. This strategy could be better applied with providing the local worker with a tele-pointing technology as well.

Most current pointing technologies for facilitating collaborative tasks are single user input tools [12,17,33,45] enabling only the remote worker to interact with the display. A single user input technology causes imbalanced communication structure between the collaborators, causing the learner to passively accept the provided information without taking it into discussion or asking for clarification. A possible technical solution to avoid unbalanced communication is to promote the local trainee's contribution to the task by enabling them to point at the view concurrently. For tasks where the trainees' hands are occupied with the task-related instruments, such as pilot training simulations and surgical training, Feng et al. [11] suggest the use of eye-gaze and voice control for pointing. The eye-gaze based pointing is also found as an effective tool in remote collaboration by Higuch et al. [23]. As such, the telepointer design may better be customized for individual tasks. While these technical enhancements may improve the communication by involving the local trainee in an active interaction, there are lessons for remote experts regarding the social context of guiding the local worker through the manipulations with significant consequences. Remote experts should be taught that simply pointing at the desired location of a significant action might not be enough for the trainee to proceed. In such cases, a verbal clarification that confirms the correctness of the pointer location is needed. Otherwise, the trainee initiates confirmation. This behavior might be even more essential in a real world work settings such as the surgical setting. In real world settings, some manipulations are more critical than others. Trainers should take this into account to flexibly use the telepointer and add verbal confirmations with regards to the significance level of the task outcomes.

Limitations

It is noteworthy that the drawing feature of the telepointer was never used by the trainer in this study. We do not have the information why since this study does not include a post-interview phase. Furthermore, this behavior could be this specific trainer's way of using this technology. Therefore, we were not able to conclude any finding out of it. However, we consider as a very interesting topic for future research. By recruiting more trainers in the future, and monitoring them with regards to the usage of a drawing function, we could obtain better insights on this tool.

6 CONCLUSION

In this study we investigated the challenges in using the information conveyed through a telepointer in remote instructional collaborative tasks by applying a conversation analysis approach to explore talk and actions between a trainer and trainee. Our analysis revealed three

main challenges that local workers face in using the telepointer instructions. We argued that the identified challenges could be addressed through intertwined social and technical implications. We further presented possible technical implications, which suggest telepointing system design improvements, and integrated social implications, which propose solutions based on training remote and local workers how to better use the telepointing technologies. In the future, the ability of presented socio-technical solutions for addressing the identified challenges should be investigated.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Grants IIS #1422671 and #1552837.

REFERENCES

- [1] Judith Amores, Xavier Benavides, and Pattie Maes. 2015. ShowMe. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. <https://doi.org/10.1145/2702613.2732927>
- [2] K. M. Augestad, H. Han, J. Paige, T. Ponsky, C. M. Schlachta, B. Dunkin, and J. Mellinger. 2017. Educational implications for surgical telementoring: a current review with recommendations for future practice, policy, and research. *Surgical Endoscopy* 31, 10: 3836–3846.
- [3] David Benyon, Phil Turner, and Susan Turner. 2005. *Designing Interactive Systems: People, Activities, Contexts, Technologies*. Pearson Education.
- [4] Eric A. Bier and Steven Freeman. 1991. MMM. *Proceedings of the 4th Annual ACM Symposium on User Interface Software and Technology - UIST '91*. <https://doi.org/10.1145/120782.120791>
- [5] S. A. Bly and S. L. Minneman. 1990. Commune: a shared drawing surface. *Proceedings of the Conference on Office Information Systems* -. <https://doi.org/10.1145/91474.91514>
- [6] Rose Challenger and Chris W. Clegg. 2011. Crowd disasters: a socio-technical systems perspective. *Contemporary Social Science* 6, 343–360. <https://doi.org/10.1080/21582041.2011.619862>
- [7] Yuan-Chia Chang, Hao-Chuan Wang, Hung-Kuo Chu, Shung-Ying Lin, and Shuo-Ping Wang. 2017. AlphaRead. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing - CSCW '17*. <https://doi.org/10.1145/2998181.2998258>
- [8] Nicklas H. Christensen, Oliver G. Hjermitsev, Frederik Falk, Marco B. Madsen, Frederik H. Østergaard, Martin Kibsgaard, Martin Kraus, Johan Poulsen, and Jane Petersson. 2017. Depth cues in augmented reality for training of robot-assisted minimally invasive surgery. *Proceedings of the 21st International Academic Mindtrek Conference on - AcademicMindtrek '17*. <https://doi.org/10.1145/3131085.3131123>
- [9] Dragoş Dăţcu, Stephan G. Lukosch, and Heide K. Lukosch. 2016. Handheld Augmented Reality for Distributed Collaborative Crime Scene Investigation. *Proceedings of the 19th International Conference on Supporting Group Work - GROUPE '16*. <https://doi.org/10.1145/2957276.2957302>
- [10] Jeff Dyck, Carl Gutwin, Sriram Subramanian, and Christopher Fedak. 2004. High-performance telepointers. *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work - CSCW '04*. <https://doi.org/10.1145/1031607.1031636>
- [11] Yuanyuan Feng, Katie Li, Azin Semsar, Hannah McGowan, Jacqueline Mun, H. Reza Zahiri, Ivan George, Adrian Park, Andrea Kleinsmith, and Helena M. Mentis. 2019. Communication Cost of Single-user Gesturing Tool in Laparoscopic Surgical Training. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. <https://doi.org/10.1145/3290605.3300841>
- [12] Yuanyuan Feng, Hannah McGowan, Azin Semsar, Hamid R. Zahiri, Ivan M. George, Timothy Turner, Adrian Park, Andrea Kleinsmith, and Helena M. Mentis. 2018. A Virtual Pointer to Support the Adoption of Professional Vision in Laparoscopic Training. *International Journal of Computer Assisted Radiology and Surgery*. 13.9 (2018): 1463-1472.
- [13] Susan R. Fussell, Robert E. Kraut, and Jane Siegel. 2000. Coordination of Communication: Effects of Shared Visual Context on Collaborative Work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW '00)*, 21–30.
- [14] Susan R. Fussell, Leslie D. Setlock, Elizabeth M. Parker, and Jie Yang. 2003. Assessing the value of a cursor pointing device for remote collaboration on physical tasks. *CHI '03 Extended Abstracts on Human Factors in Computing Systems - CHI '03*. <https://doi.org/10.1145/765891.765992>
- [15] Susan R. Fussell, Leslie D. Setlock, Jie Yang, Jiazi Ou, Elizabeth Mauer, and Adam D. I. Kramer. 2004. Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks. *Human-Computer Interaction* 19, 273–309. https://doi.org/10.1207/s15327051hci1903_3
- [16] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. In touch with the remote world. *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14*. <https://doi.org/10.1145/2671015.2671016>
- [17] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. 2014. World-stabilized annotations and virtual scene navigation for remote collaboration. *Proceedings of the 27th annual ACM Symposium on User Interface Software and Technology - UIST '14*. <https://doi.org/10.1145/2642918.2647372>
- [18] Pavel Gurevich, Joel Lanir, and Benjamin Cohen. 2015. Design and Implementation of TeleAdvisor: a Projection-PACM on Human-Computer Interaction, Vol. 3, No. CSCW, Article 93, Publication date: November 2019.

- Based Augmented Reality System for Remote Collaboration. *Computer Supported Cooperative Work (CSCW)* 24, 527–562. <https://doi.org/10.1007/s10606-015-9232-7>
- [19] Carl Gutwin and Reagan Penner. 2002. Improving interpretation of remote gestures with telepointer traces. *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work - CSCW '02*. <https://doi.org/10.1145/587078.587086>
- [20] Christian Heath, Jon Hindmarsh, and Paul Luff. 2010. Video in Qualitative Research: Analysing Social Interaction in Everyday Life. <https://doi.org/10.4135/9781526435385>
- [21] Christian Heath, Paul Luff, Hideaki Kuzuoka, Keiichi Yamazaki, and Shinya Oyama. Creating Coherent Environments for Collaboration. *ECSCW 2001*, 119–138. https://doi.org/10.1007/0-306-48019-0_7
- [22] Alexa Hepburn and Galina B. Bolden. 2012. The Conversation Analytic Approach to Transcription. *The Handbook of Conversation Analysis*, 57–76. <https://doi.org/10.1002/9781118325001.ch4>
- [23] Keita Higuch, Ryo Yonetani, and Yoichi Sato. 2016. Can Eye Help You? *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*. <https://doi.org/10.1145/2858036.2858438>
- [24] Jon Hindmarsh and Christian Heath. 2000. Embodied reference: A study of deixis in workplace interaction. *Journal of Pragmatics* 32, 1855–1878. [https://doi.org/10.1016/s0378-2166\(99\)00122-8](https://doi.org/10.1016/s0378-2166(99)00122-8)
- [25] Weidong Huang, Leila Alem, Franco Tecchia, and Henry Been-Lirn Duh. 2018. Augmented 3D hands: a gesture-based mixed reality system for distributed collaboration. *Journal on Multimodal User Interfaces* 12, 77–89. <https://doi.org/10.1007/s12193-017-0250-2>
- [26] Hiroshi Ishii and Minoru Kobayashi. 1992. ClearBoard. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '92*. <https://doi.org/10.1145/142750.142977>
- [27] Rohana Abdul Karim, Nor Farizan Zakaria, Mohd Asyraf Zulkifley, Mohd Marzuki Mustafa, Ismail Sagap, and Nani Harlina Md Latar. 2013. Telepointer technology in telemedicine : a review. *BioMedical Engineering OnLine* 12, 21. <https://doi.org/10.1186/1475-925x-12-21>
- [28] Laurent Karsenty. 1999. Cooperative Work and Shared Visual Context: An Empirical Study of Comprehension Problems in Side-by-Side and Remote Help Dialogues. *Human-Computer Interaction* 14, 3: 283–315.
- [29] Seungwon Kim, Gun A. Lee, and Nobuchika Sakata. 2013. Comparing pointing and drawing for remote collaboration. *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. <https://doi.org/10.1109/ismar.2013.6671833>
- [30] David Kirk, Tom Rodden, and Danaë Stanton Fraser. 2007. Turn it this way. *Proceedings of the SIGCHI conference on Human Factors in Computing Systems - CHI '07*. <https://doi.org/10.1145/1240624.1240782>
- [31] David S. Kirk and Danaë Stanton Fraser. 2005. The effects of remote gesturing on distance instruction. *Proceedings of the 2005 Conference on Computer Support for Collaborative Learning 2005: the next 10 years! - CSCW '05*. <https://doi.org/10.3115/1149293.1149332>
- [32] Sotaro Kita. 2003. *Pointing: Where Language, Culture, and Cognition Meet*. Psychology Press.
- [33] Hideaki Kuzuoka, Toshio Kosuge, and Masatomo Tanaka. 1994. GestureCam. *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work - CSCW '94*. <https://doi.org/10.1145/192844.192866>
- [34] Hideaki Kuzuoka, Jun Yamashita, Keiichi Yamazaki, and Akiko Yamazaki. 1999. Agora. *CHI '99 extended abstracts on Human Factors in Computing Systems - CHI '99*. <https://doi.org/10.1145/632835.632836>
- [35] Jiann-Shu Lee, Ching-Tsorng Tsai, Chen-Hsing Pen, and Hui-Chieh Lu. 2003. A real time collaboration system for teleradiology consultation. *International Journal of Medical Informatics* 72, 1-3: 73–79.
- [36] Paul Luff, Christian Heath, Hideaki Kuzuoka, Keiichi Yamazaki, and Jun Yamashita. 2006. Handling documents and discriminating objects in hybrid spaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '06*. <https://doi.org/10.1145/1124772.1124858>
- [37] Helena M. Mentis, Amine Chellali, and Steven Schwaitzberg. 2014. Learning to see the body. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*. <https://doi.org/10.1145/2556288.2557387>
- [38] Helena M. Mentis, Kenton O'Hara, Gerardo Gonzalez, Abigail Sellen, Robert Corish, Antonio Criminisi, Rikin Trivedi, and Pierre Theodore. 2015. Voice or Gesture in the Operating Room. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*. <https://doi.org/10.1145/2702613.2702963>
- [39] Helena M. Mentis, Ahmed Rahim, and Pierre Theodore. 2016. Crafting the Image in Surgical Telemedicine. *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16*. <https://doi.org/10.1145/2818048.2819978>
- [40] Microsoft Corporation. Remote Assist | Microsoft Dynamics 365. Retrieved March 31, 2019 from <https://dynamics.microsoft.com/en-us/mixed-reality/remote-assist/>
- [41] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2015. Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. <https://doi.org/10.1145/2807442.2807497>
- [42] Kenton O'Hara, Gerardo Gonzalez, Graeme Penney, Abigail Sellen, Robert Corish, Helena Mentis, Andreas Varnavas, Antonio Criminisi, Mark Rouncefield, Neville Dastur, and Tom Carrell. 2014. Interactional Order and Constructed Ways of Seeing with Touchless Imaging Systems in Surgery. *Computer Supported Cooperative Work: CSCW: an International Journal* 23, 3: 299–337.
- [43] Shoichi Ohta, Hideaki Kuzuoka, Mariko Noda, Hirokazu Sasaki, Shiro Mishima, Tadashi Fujikawa, and Tetsuo Yukioka. 2006. Remote support for emergency medicine using a remote-control laser pointer. *Journal of Telemedicine and Telecare* 12, 1: 44–48.

- [44] Noritaka Osawa. 2006. Aggregate pointers to support large group collaboration using telepointers. *CHI '06 Extended Abstracts on Human Factors in Computing Systems - CHI EA '06*. <https://doi.org/10.1145/1125451.1125671>
- [45] Jiazhi Ou, Susan R. Fussell, Xilin Chen, Leslie D. Setlock, and Jie Yang. 2003. Gestural communication over video stream. *Proceedings of the 5th International Conference on Multimodal Interfaces - ICMI '03*. <https://doi.org/10.1145/958468.958477>
- [46] Todd A. Ponsky, Iuliana D. Bobanga, Marc Schwachter, Theodore H. Stathos, Michael Rosen, Robert Parry, Margaret Nalugo, and Steven S. Rothenberg. 2014. Transcontinental telementoring with pediatric surgeons: proof of concept and technical considerations. *Journal of Laparoscopic & Advanced Surgical Techniques. Part A* 24, 12: 892–896.
- [47] Steven S. Rothenberg, Suzanne Yoder, Sandra Kay, and Todd Ponsky. 2009. Initial experience with surgical telementoring in pediatric laparoscopic surgery using remote presence technology. *Journal of Laparoscopic & Advanced Surgical Techniques. Part A* 19 Suppl 1: S219–22.
- [48] Christopher M. Schlachta, Ninh T. Nguyen, Todd Ponsky, and Brian Dunkin. 2016. Project 6 Summit: SAGES telementoring initiative. *Surgical endoscopy* 30, 9: 3665–3672.
- [49] Azin Semsar, Hannah McGowan, Yuanyuan Feng, Hamid R. Zahiri, Ivan M. George, Timothy Turner, Adrian Park, Helena M. Mentis, and Andrea Kleinsmith. 2019. Effects of a Virtual Pointer on Trainees' Cognitive Load and Communication Efficiency in Surgical Training. *AMIA Annual Symposium Proceedings 2019*. American Medical Informatics Association.
- [50] Duncan Roderick Stevenson. 2011. Tertiary-Level Telehealth: A Media Space Application. *Computer Supported Cooperative Work (CSCW)* 20, 61–92. <https://doi.org/10.1007/s10606-010-9125-8>
- [51] Barry A. Turner. 1978. *Man-made Disasters*. Crane Russak, Incorporated.
- [52] Nirit Yuviler-Gavish, Eldad Yechiam, and Arava Kallai. 2011. Learning in multimodal training: Visual guidance can be both appealing and disadvantageous in spatial tasks. *International Journal of Human-Computer Studies* 69, 113–122. <https://doi.org/10.1016/j.ijhcs.2010.11.005>
- [53] 2019. New AI, mixed reality business solutions lead the way for Microsoft Dynamics 365 - Microsoft Dynamics 365. *Microsoft Dynamics 365*. Retrieved March 31, 2019 from <https://cloudblogs.microsoft.com/dynamics365/2019/02/21/new-ai-mixed-reality-business-solutions-lead-the-way-for-microsoft-dynamics-365/>

Received April 2019; revised June 2019; accepted August 2019.