Measuring the quality of learning in a human-robot collaboration: a study of laparoscopic surgery
Eléonore Ferrier-Barbut, Philippe Gauthier, Vanda Luengo, Geoffroy Canlorbe, Marie-Aude Vitrani

To cite this version:

HAL Id: hal-03355055
https://hal.science/hal-03355055
Submitted on 8 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ELÉONORE FERRIER-BARBUT, Sorbonne Université, France
PHILIPPE GAUTHIER, Sorbonne Université, France
VANDA LUENGO, Sorbonne Université, France
GEOFFROY CANLORBE, Sorbonne Université, France
MARIE-AUDE VITRANI, Sorbonne Université, France

Robot-Assisted Laparoscopic Surgery (RALS) is now prevalent in Operating Rooms (ORs). This situation requires future surgeons to learn Classic Laparoscopic Surgery (CLS) and RALS simultaneously. Therefore, along with the investigation of the differences in performance between the two techniques, it is essential to study the impact of training in RALS on the skills mastered in CLS. In this article, we study comanipulated RALS (Co-RALS), one of the two designs for RALS, where the human and the robot share the execution of the task. We use a rarely used in Human-Robot Interaction measuring tool: gaze tracking, and time recording to measure for the acquisition of skills in CLS either when training in Co-RALS or in CLS, and time recording to compare the learning curves between Co-RALS and CLS. These metrics allow us to observe differences in Co-RALS and CLS. Training in Co-RALS develops slightly but not significantly better hand-eye coordination skills and significantly better time-wise performance compared with training in CLS alone. Co-RALS enhances time-wise performance in laparoscopic surgery on specific type of task that requires precision rather than depth perception skills, compared with CLS. The results obtained enable to further define the Human Robot Interaction quality in Co-RALS.

CCS Concepts: • Human-centered computing → Collaborative interaction; Empirical studies in interaction design.

ACM Reference Format:

1 INTRODUCTION

Robotic surgery is increasingly used in Operating Rooms (ORs), especially the Da Vinci (cf. figure 1-j) with over 1 million procedures performed in 2018 [36]. This robot is essentially used to facilitate the practice of laparoscopic surgery. Without robot, e.g. Classic Laparoscopic Surgery (CLS) (cf. figure 1-e) is surgery done through small holes in the patient’s belly, in which are inserted long instruments and a camera i.e. endoscope. The image of the endoscope is displayed on a 2D screen, showing the inside of the patient’s body. CLS is advantageous for patients (less pain, shorter recovery, aesthetic benefits) but disadvantageous for surgeons who then have more visual as well as physical difficulties to deal with than in open surgery. Visual difficulties are partly generated by the absence of direct vision. The working space is visualized on a 2-dimensional screen while gestures are performed in a 3-dimensional space,
making hand-eye coordination and depth perception complicated [7]. Physical difficulties are due to the length of the instruments and to the fact that their insertion point (the small incisions in the patient body) induces kinematic restrictions and the appearance of the fulcrum or lever arm effect [25]. Robot-Assisted Laparoscopic Surgery (RALS) aims to alleviate varying difficulties depending on the robot’s interface and features. There are two interaction designs for Robot Assisted Laparoscopic Surgery (RALS): telemanipulated (Tele-RALS) (cf. figure 1-f) such as with the Da Vinci and comanipulated (Co-RALS) (cf. figure 1-k). In Tele-RALS, as indicated by the semantics, a distance exists between the surgeon (who sits at the master console) and the patient (above who stand the robotic arms holding the instruments, the slave console). This distance removes the hand-eye coordination and the physical difficulties as instruments situated at the slave console are controlled with joysticks situated at the master’s console thanks to a communication channel between the two. However, it introduces new difficulties. To use the robot surgeons and future surgeons have to learn, on top of traditional skills linked with Classic Laparoscopic Surgery (CLS) more technology-oriented skills. Here, we study the learning period when surgeons simultaneously learn to perform classic (without robot) and robot-assisted laparoscopic surgery. During this period, the discrepancy between psycho-motor skills required to perform the two techniques is problematic. Most importantly since learning laparoscopic surgery with a telemanipulated assistance does not result in a transfer of skills to the classic technique [4], a situation that requires re-learning for each switch between techniques.

A comanipulated robotic assistance, such as the Acrobot [17], aims to improve security and dexterity when performing the surgical gesture while keeping the surgeon close to her/his patient. Comanipulation means the shared control of the instruments by the robotic assistance and the surgeon [24] (cf. figure 1-l): both the robotic arms and the user hold the instruments needed to perform the surgery. Same as for laparoscopic surgery, while technically possible to perform viewing the image of the endoscope in 3D, Co-RALS is most frequently performed viewing the image of the endoscope on a 2D screen. This design is at the fourth degree of human-robot collaboration scale [14]: Supportive. At this level, the human and the robot work together at the same time and with the same workpiece to complete a common task. Arguably, a successful Supportive human-robot collaboration is defined by its ability to augment performance compared to the same task performed without a robot, but also to confer to the human a cognitive and physical load that is neither too important nor too little, to diminish the difficulty of the task while keeping the human involved and active [27]. For a task as complex as surgery simultaneously performed by a human operator and a robot, where robots are far from being able to replace the human operators, it can be suggested that keeping the human involved and active is mandatory, for him/her to maintain his/her acquired skills level. Indeed, in our case, the robot assists the human in a task s/he already masters. Comanipulated robots for surgery are destined to work very closely with surgeons and future surgeons without disturbing their dexterity skills and while increasing their performance. The robot only augments already existing technical skills. Still, the quality of the human-robot interaction in Co-RALS has not been investigated yet to our knowledge. We propose a method to measure with quantitative metrics, how a collaborative robotic assistance not only can help to increase performance and learning in very complex tasks but also the impact it has on the technical skills mastered by the human. We study it through two research questions, situated in the specific context of learning when the acquisition of skills is of crucial importance.

The main research question is the following: does training with a comanipulated robot in what we call a target task, in our case exercises of laparoscopic surgery, develop hand-eye coordination and time-wise skills in this target -and simultaneously learned in a surgical curriculum- task? We focus on the hand-eye coordination skills as these are especially complex to learn and perform in laparoscopic surgery for the previously mentioned reasons. To measure the hand-eye coordination skills, we use a tool that to our knowledge has rarely been used in Human-Robot Interaction:
gaze tracking. The number of fixations (the event when the gaze remains on a point for 50 to 600 ms [18]) on the aimed target, the fixation rate per second and the duration of fixations, when performing in CLS are measured either after training in Co-RALS or in CLS. Previous works have shown that these metrics enable to measure hand-eye coordination skills developed when training [13], in laparoscopic surgery especially [19, 42, 43]. To measure time-wise skills, time elapsed is recorded when performing exercises in CLS, either after training in Co-RALS or in CLS. This shows the operator’s performance efficiency [22] in the target task after training in Co-RALS. The secondary research question is the following: does using Co-RALS improve time-wise performance on exercises of laparoscopic surgery compared with the same exercises performed in CLS? Time elapsed is also recorded when performing repeated exercises with each technique. The repetition of the exercises highlights the shapes of the learning curves. Other metrics are used as exploratory measures: the NASA TLX [12] that serves to compare the workload between the robot-assisted technique and the classic technique, and a performance score given for the exercise performed in the classic technique either after training with the robotic-assistance or without.
In section 2 the existing connections between motor learning, robotic systems and their interfaces are discussed and comanipulation is presented. In section 3 the experiment’s methods are detailed. In section 4 results are submitted. The results presented below, as measured by gaze tracking and time recording suggest that Co-RALS, contrarily to what was shown for Tele-RALS, does not negatively impact skills in CLS, developing slightly but not significantly better hand-eye coordination skills. Additionally, during the training session, the exercise of laparoscopic surgery that develops precision skills presents a significantly shorter learning curve when performed in Co-RALS compared with CLS, while an exercise of laparoscopic surgery that develops depth perception skills presents a learning curve that is not statistically different from that of CLS. The results, although reporting on a relatively small number of participants, demonstrate the ability of the metrics chosen to show differences in development of hand-eye coordination skills between the classic and the robot-assisted techniques for laparoscopic surgery, and differences in the learning curve between the two.

2 STATE OF THE ART

In this section, the consequences of interaction design for RALS on skills mastered by the user are presented and comanipulation is described.

2.1 Performance, human skills, and interfaces for RALS

Numerous articles in the field of Human Computer Interfaces (HCI) have focused on the impact of interfaces for RALS on non-technical skills [45] related to laparoscopic surgery: workflow, communication, situation awareness, teamwork [2, 6, 29, 32–34]. These works are exclusively on Tele-RALS, and the metrics used are essentially subjective. These robotic systems enhance the surgeon’s experience by making them more independent to perform surgery and increasing the number of instruments they can control. But the distance imposed with patients keeps them away from their team. This physical distance is compounded by visual, auditory and mental distances. It disrupts access to information, changes power distribution and decreases the surgeon’s situation awareness. It also has consequences for students who then have fewer tasks, decisions and actions to make [41] than when the surgical intervention is performed without the robot. Their learning process is made more difficult: they can no longer learn by observing, hearing, and doing according to the well known wording See one, do one, teach one [30] as they are not standing next to the surgeon any more [23].

Although less studied in HCI, the impact of interaction design for robotic assistance on technical skills [26] related to laparoscopic surgery is just as strong. The interaction design for robotic assistance defines the nature and the difficulty of motor skills either learned or mastered by the user, and how they transfer to without robotic assistance. Robotic assistance can easily improve performance when being used. However, as stressed in a review of literature on robotic assistance of motor learning by Heuer et. al’s from 2015 from the field of neurosciences, it is complicated to ensure that after the robot is turned off, the user continues to perform as well as if s/he had never used it [15]. A risk exists that the use of robotic assistance becomes “normal” and the motor learning highly dependent on the robot specific dynamic environment. This situation would lead to a negative impact of training with a robotic assistance on skills mastered in the classic technique, as it seems to be the case for Tele-RALS [3, 4]. Both these studies by Blavier et al. from 2007 were conducted with medical students without any prior surgical experience and results suggest that training in Tele-RALS has negative consequences on mastery of skills in CLS, compared with training in CLS alone. In the case of a conversion from Tele-RALS to CLS during a surgical intervention, a scenario that regularly happens for different reasons such as mentioned in Blavier’s et al. article from 2007 [3], this negative impact can have major consequences.
Here, we hypothesize that training in Co-RALS results in equivalent hand-eye coordination and time-wise abilities compared with training in CLS. Hence we take a different approach than previous work: with objective metrics, we seek to observe the quality of the human-robot interaction through the analysis of the skills developed by the human in conjunction with the performances achieved by the human-robot team.

### 2.2 Comanipulated Interfaces in Robotics

The definition of comanipulation can be found in Morel’s article from 2013: A comanipulator is thus any robotic system performing a task, most often in contact with the environment, that can be controlled through direct contact by an operator. It aims to increase the manipulation performance of the operator [24].

Comanipulation is an interaction paradigm involving a robot and a user simultaneously manipulating a load or a tool (cf. figure 1-m). The robot is employed as a comanipulated device, in the sense that the gesture control of the instrument is shared by the robot and the surgeon. Of the difficulties of CLS mentioned earlier, comanipulation aims to facilitate the physical ones in particular. The currently existing commercialized comanipulated robotic systems are basically designed for specific types of surgical tasks [10, 16, 17, 28]. Research institutes have also exploited the idea for precise surgical tasks [5]. Comanipulation can be applied to tasks that require both precise manipulation and human judgment so as to enhance gesture quality [37]. This interaction design for robotics has the ability to at least compensate for gravity and filter tremors [31] while performing the surgical gesture. Some other physical difficulties can also be dealt with including the reduction of the fulcrum effect for example with the help of an active force feedback [35].

To improve haptic feedback, virtual fixtures can be applied [11, 40]. In this study, two comanipulated robotic arms are used. One for each instrument manipulated when performing exercises in laparoscopic surgery. Only the basic functions of the robots are studied. Gravity compensation for instruments is ensured, as well as tremor filtration. The algorithm used for tremor filtration is a viscous field, i.e. a damping algorithm proportional to the velocity of movement [8]. The robots exhibit high viscosity at low velocities and no viscosity at high velocities, at the instrument tip. It helps the surgeons during precise surgical tasks (performed at small speed), by filtering unintentional movements and augmenting precision of the gesture. This viscous field also increases the rate of adaptation of the user, forced to remain active when performing her/his task. A similar algorithm, implemented in a comanipulated robotic assistance where the human controls the direction and the speed while the robot ensures the precision and smoothness of motions by suppressing sudden and abrupt gestures, has been shown to significantly improve performance for tasks of manual welding [9]. In an other research article [39], a human robot cooperative calligraphic task is performed, in which the human and the robot grasp a writing brush at the same time. The robot is controlled against the human force to prevent the vibration and to enhance the accuracy, and results show the advantage of the control method. The benefits of variable damping coefficient have been demonstrated. Thus, referring to the second question of research, it is hypothesized that in Co-RALS, the adaptive damping algorithm implemented enables better time-wise performance for exercises of laparoscopic surgery compared to the same exercises performed without the robotic assistance.

### 3 METHODS

To investigate 1. whether training with a comanipulated device in laparoscopic surgery results in an equivalent development of skills in this target task, exercises of laparoscopic surgery, compared with training in the target task and 2. the comparison of the learning curve between the robot-assisted task and the classic task, we present an experimental
protocol involving two conditions. The participants in Classic condition are pre-tested in CLS during the Pre-Learning Step, then trained in CLS, and then post-tested during the Post-Learning Step in CLS. The participants in Robot condition are also pre-tested in CLS, then trained in RALS, then post-tested in CLS (cf. table 1). Both groups are Pre-trained and Post-trained in the same task, CLS, to compare the differences in the process of learning for each group, Classic and Robot, in CLS. The main research question is studied by comparing the skills in CLS of the two groups, Classic and Robot with gaze-tracking and time recording, and the secondary research question by comparing the learning curves of the two groups Classic and Robot during their learning session with time recording. The exercises of laparoscopic surgery performed at each step were chosen among basic training exercises for this discipline, depending on the skills they enable to train and measure. During the Pre and Post-Learning steps, the exercise chosen was the Peg Transfer as it enables to observe a large panel of participant’s skills in laparoscopic surgery: bimanual coordination, precision, and depth perception. It requires the participant to lift six objects place on the left side of a board with a grasper first using non-dominant (in this case, left) hand and transfer the object midair to the dominant hand. Then, the participant has to place each object on a peg on the right side of the board. It is also performed from dominant to non-dominant hand. This exercise is further described later on. Pea on a Peg and Loops and Wire were chosen because they trained precision and depth perception respectively. Exercise Pea on a Peg consists of placing 14 beads on pegs of different highs placed on a board, and exercise Loops and Wire consists of passing a wire through 4 different loops placed on a board. Both these exercises are also further described later on. The fact that each exercise trains different skills enables to better define the differences between CLS and Co-RALS in terms of learning. This experimental protocol is tested both with Resident participants and, more numerous, Non-Resident participants. The similarity of the exercises in laparoscopic surgery performed by each group is controlled. Level equivalence between group Robot and Classic is statistically verified for Non-Resident participants during the Pre-Learning Step. There were not enough Resident participants involved in the protocol to perform this verification, hence the Resident participants’ results are presented as indication of tendency and insights for future research rather than results that can be generalized. The experiment is thus divided into three steps: 1. A pre-learning step to control the participants’ level, 2. A learning step to train the participants of each group respectively in each task and (CLS and Co-RALS) and compare their performance and 3. A post-learning step to compare the mean level of each group in the target activity, i.e. CLS.

3.1 Participants

Fourteen Non-Resident participants and six Residents in medicine are recruited. Among them, seven Non-Resident participants and three Resident participants (group Classic) performed their training session in in CLS (task CLS). Seven other Non-Resident participants and three other Resident participants (group Robot) performed their training
session in RALS on a comanipulated robot (task Co-RALS). Resident participants have a small and contrasted experience in laparoscopic surgery ($6 \pm 4.5$ hours of practice). The number of Resident participants involved in the study does not allow to perform statistics. Still, their results enable to verify whether the trend observed among novices may be applied to intermediate level users. Because of their tight agenda, Residents in medicine could only perform one exercise, *Pea on a Peg*, during the training session and, due to a lack of technical and time resources, no gaze data could be recorded during *Peg Transfer* exercise before and after the Learning session. Their results are presented separately. Five of them are first year residents and one is third year resident. Non-Resident participants all declared to be novices in laparoscopic surgery, and are all university students. All procedures are in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

### 3.2 Material

3.2.1 *Apparatus.* Material includes two set-ups: one in CLS, and one in Co-RALS.

**Task CLS** The CLS set-up: It includes two surgical graspers. Exercises are performed in a pelvi-trainer, above which is a 2D screen transcribing the 3D working space inside it (*cf. figure 2a*). The endoscope does not move during the entire learning session.

**Task Co-RALS** The Co-RALS set-up: It consists of two robotic arms, each holding the same two graspers as those used in the CLS set-up (*cf. figure 2b*). The two robotic arms are modified Haption Virtuose 3D robots, characterized by six rotational joints. The robots used are 3D robots as other degrees of freedom are constrained by the entree point in the patients’ belly. The first three joints are fully actuated and the other three form a free wrist that allows full motion across the surgical workspace [21]. Our research team designed the software implemented in these robotics arms: (i) tool weight compensation so that holding the tools is transparent to surgeons [38] and (ii) a damping algorithm proportional to the velocity of movement, experimentally tuned such as in [8, 20]. The same pelvi-trainer as for the CLS set-up is used (*cf. figure 2a*). The endoscope is not moved during the entire learning session.
3.2.2 Measuring tools and metrics. Different measuring tools and metrics are used at each step of the experiment to assess hand-eye coordination abilities, time-wise performance and, workload and scores. They were chosen for their ability to show not only the observed performance at the end of a task but also the learning process itself. Gaze-tracking is used during the Pre-Learning Step to verify the absence of heterogeneity in average hand-eye coordination skills in laparoscopic surgery of groups Classic and Robot before the Learning session. It is also used during Post-Learning step to quantify for differences in the same skills between these groups after the Learning session. Time is recorded at each step of the experiment. Before and after the Learning session this metric enables to answer the first research question, together with gaze-tracking. During the learning session, Time recording serves to answer the secondary research question. The results obtained at the NASA Task Load Index during the learning session and the scores at the Peg Transfer exercise are explored, respectively as self-perceived feeling of comfort with the task and observed performance.

Pre-Learning Step

- Gaze-Tracking: Observing gaze pattern has been proven to be relevant in distinguishing novices from experts [13, 19, 42, 44]. Experts have a greater ability to anticipate their gestures with their gaze than novices. This results, in tasks of target reaching (such as the Peg Transfer), in a greater ability to look at the aimed target rather than the instruments used to reach the target. In other words, experts anticipate their movement by taking their eyes off it and projecting them to its goal, while novices are focused on achieving the movement itself and keep their eyes on it. This translates into a struggle to detach their eyes off their hands or the instruments they hold to perform the movement and to project their gaze to its goal. It can be measured by the number of fixations (the maintenance of the gaze in a single location) on the aimed target before reaching it with the instruments, a greater number meaning a better ability to anticipate, and greater expertise [1]. Experts also tend to do less back and forth movements with the eyes -movements used by novices to make calculations of distances, lengths, velocities- and a smaller number of fixations. The duration of their fixations also tends to be larger than for novices. Hence, the number of fixations on the aimed target before it has been reached with the instruments, fixation rate per second and duration of fixations are analyzed. A higher number of fixations on the aimed target at the specific moment of reaching for the target in the Peg Transfer means a greater level of comfort with the task. At the opposite, less fixations per second and greater duration of fixations while performing the entire exercise means better hand-eye coordination skills. Gaze data analysis is performed separately for each phase of the Peg Transfer exercise described later on in the article. It is cut into 3 phases: Grab, Transfer and Drop. Grab corresponds to the moment where participants grab the bead with their dominant hand, Transfer to the moment they pass it to their non-dominant hand and Drop to the moment they put it on the peg. The Transfer and Drop phases are the most interesting. They represent the moment where the participants are specifically doing a target reaching task. Recording gaze data enables observation of hand-eye coordination abilities. The Eye Tracker used to record gaze data is a Tobii X3-120 screen-based eye tracker with a sampling rate of 120Hz. It is mounted at the bottom of the screen at the eye level of participants. Data analysis is performed with Tobii Pro Lab and the filtering algorithm used is the IV-T Classifier. To identify the number of fixations on the aimed target, we manually defined, using Tobii Pro Lab, dynamic zones of interest. They were the zones, on the 2D screen, aimed by the participants when performing the Peg Transfer exercise described later on. Statistics of the number of fixations in these zones can then be extracted.
- Time: A timer is set during performance of exercises.
Learning: from robotic assistance to without. Woodstock ’18, June 03–05, 2018, Woodstock, NY

Fig. 3. (a) Peg Transfer (b) Pea on a Peg (c) Loops and Wire

- Score: Videos of the camera used to perform the exercises is recorded to allow for scoring of the Peg Transfer exercise. 10 penalty points are counted per dropped sleeve. When a sleeve falls from the pegboard 20 penalty points are counted. Total score is time in seconds + penalty points. This score shows temporal as well as qualitative performance.

Learning Step
- Time: A timer is set during each trial of the two exercises performed during the learning session, described latter on in the article.
- NASA Task Load Index (TLX): It is used as an exploratory measure of comfort level with the executed tasks during the learning session. This index is a subjective, multidimensional assessment tool that rates perceived workload in order to assess a task, system, or team’s effectiveness or other aspects of performance. It was developed by the Human Performance Group at NASA’s Arms Research Center. The higher the score is, the more important is the perceived workload. In this study, the score is NASA weighted. Participants were required to fill the NASA TLX after they had performed the learning session.

Post-Learning Step
The measuring tools and metrics are the same as during the Pre-Learning Step.

3.3 Procedures
Experiment is divided into three steps (cf. table 1):

3.3.1 Pre-Learning Step. This step consists in the measurement of participants’ base level. The data recorded enables to ensure that the two groups of participants: Classic and Robot do not show, initially, significant differences in level in the target activity. Means of each group before the learning session are compared to check for statistical differences. The exercises performed are the following:

- **Peg Transfer Non Dominant:** familiarization exercise, data is not recorded during this task because it is the familiarization task, during which participants discover instruments and set-up for four minutes. There are 6 sleeves and 12 pegs on the board. The 6 sleeves are positioned on 6 pegs, on one side of the board. Subjects have to transfer the sleeves from one side of the board to the other. To pick up a sleeve, they use the instrument held in the non dominant hand. Then they transfer the sleeve to the instrument held in the dominant hand. Maximum time is 4 minutes.
• **Peg Transfer Dominant**: recorded exercise. Subjects have to transfer sleeves resting on pegs on the other side of the board. They pick the sleeves up with the instrument held in the dominant hand, transfer them in the instrument held in the non dominant hand, drop them on the other side of the board. They have to do as many as possible within four minutes. Data is recorded during this exercise only (cf. figure 3a).

### 3.3.2 Learning Step.
Participants are randomly assigned either to group Classic which will perform one hour training session in CLS or group Robot which will perform 1h training in Co-RALS. During each learning session, participants have to perform two exercises: *Pea on a Peg, Loops and Wire* (cf. figure 3). These two exercises are performed in a random order. They are each performed repeatedly: participants had a maximum number of five trials for each exercise. To avoid the influence of fatigue on performance, a maximum total time per exercise and a maximum time per trial is set. The exercises are the same when performed either in Co-RALS or in CLS. Exercises performed are the following:

- **Pea on a Peg**: Participants have to place 14 beads on pegs of different heights. The beads are positioned on a peg board with the cup containing them in front. A maximum time per trial of 10 minutes is set as well as a maximum number of trials of 5 and a maximum time spent on exercise of 30 minutes. Once all the beads are placed on the pegs, the trial is finished. The psychomotor skills developed are similar to those involved during *Peg Transfer*: fine motor skills, coordination, precision and depth perception (cf. figure 3b).
- **Loops and wire**: The exercise contains a peg board with 4 loops on which is positioned a flexible wire. Participants have to insert the wire in the 4 loops in a specific order indicated on the board. A maximum time per trial of 4 minutes is set as well as a maximum number of trials of 5 and a maximum of time spent on exercise of 15 minutes. Once the wire is inserted in all the loops, the trial is finished. The psychomotor skills developed are related to depth perception and manipulation of the instruments (cf. figure 3c).

### 3.3.3 Post-Learning Step.
Participants have to perform the same exercises as the one performed during the pre-learning step:

- **Peg Transfer Non Dominant**: familiarization exercise.
- **Peg Transfer Dominant**: recorded exercise.

### 4 RESULTS
We present the results given by each of the measuring tools and metrics used during the experiment. As mentioned before, gaze-tracking could only be recorded for Non-Resident participants while time and score were recorded for all participants. Statistical tests were performed as a comparison between the means of each group: *Classic* and *Robot* for every metric used during Pre-Learning, Learning and Post-Learning steps. All sets of data were normally distributed, variances were homogeneous and samples were independent, hence Student’s t tests were performed for every pair of data: number of fixations, duration of fixations, time to perform *Peg Transfer*.

#### 4.1 Main Research Question: Does learning laparoscopic surgery with a comanipulated robotic assistance result in development of skills in CLS?

##### 4.1.1 Pre-Learning Step.
Gaze data and time data during Pre-Learning is analyzed to compare between group Classic and group Robot and make sure that there are no statistical differences between these two groups previous to the learning session. The means in terms of time taken to perform *Peg Transfer* in the Pre-Learning step are different between group Classic (3.8 ± 0.4) and group Robot (2.7 ± 0.9) but not significantly different (Student’s t Test, t(11.2)=2.09,
p = 0.06). Still, because of the low p-value, the difference in terms of time taken to perform Peg Transfer after learning between group Classic and group Robot should be considered with caution. All other data measured before learning shows clearly no statistical difference between group Classic and group Robot in terms of gaze data i.e. duration of fixations during Grab phase (Student’s t Test, t(8.7459) = 0.62994, p = 0.5), Transfer phase (Student’s t Test, t(9.4918) = -0.5219, p = 0.6) and Drop phase (Student’s t Test, t(9.8688) = 0.54009, p = 0.6011).

These tests have been performed for Non-Resident participants, Resident participants were not numerous enough to perform statistical comparisons.

4.1.2 Post-Learning Step.

- Mean number of fixations on aimed target: for Non-Resident participants, before learning, on average participants perform 0.6 ± 0.1 fixations on aimed target during Transfer phase. In the same phase, after learning, group Classic performs 0.55 ± 0.2 fixations on average on aimed targets, and group Robot 0.59 ± 0.2 fixations on average on aimed target. Before learning, on average participants perform 0.54 ± 0.19 fixations on aimed target during Drop phase. In the same phase, after learning, group Classic performs 0.55 ± 0.22 fixations and Group Robot performs 0.45 ± 0.13 fixations on average on aimed target (cf. figure 4). These differences are not significant (Student’s t Test, t(0.12) = 20.3, p = 0.9).

![Fig. 4. Number of fixations on target before reaching them with the instrument during Transfer and Drop phases: Non-Residents. Ta1-3 (Mean number of fixations on Targets n°1, 2 and 3 before reaching them with the instrument), Ta4-5 (Mean number of fixations on Targets n°4 and 5 before reaching them with the instrument), Ta6 (Mean number of fixations on Target n°6 before reaching it with the instrument). Targets are separated in accordance with the % of participants who managed to reach them, this number being proportionally lower for the targets with the higher number. On average, group Robot does more fixations on targets 1 to 5 in Transfer phase, on targets 1 to 3 in Drop phase, than group Classic. Group Classic does more fixations than group Robot on target n°6, but variability is also twice as important: ±0.28 for group Classic, ±0.14 for group Robot.](image-url)
• Mean fixation rate per second: for Non-Resident participants, the fixation rate per second decreases for both groups before and after learning all phases, except in Grab phase where it slightly increases for group Classic. After learning, in Grab phase, group Classic increases on average the fixation rate per second of 4% compared to before learning, group Robot stagnates. In Transfer phase, group Classic does on average as many fixations per second before and after learning, while group Robot does 5% less. Finally in Drop phase, group Classic decreases by 6% on average the fixation rate per second, group Robot decreases by 10% (cf. figure 5). Difference is not statistically significant between the groups (Student’s t Test, t(1.9)=9.7, p = 0.08).

• Duration of fixations: for Non-Resident participants, the duration of fixations increases for both groups before and after learning in all phases of the Peg Transfer exercise, except for Grab phase in which it slightly decreases for group Classic. After learning, in Grab phase, group Classic decreases on average the duration of fixations by 9% compared to before learning, group Robot increases by 6%. In Transfer phase, group Classic increases the duration of fixations of 3% before and after learning, while group Robot increases by 5%. Finally in Drop phase, group Classic increases by 2% on average the duration of fixations, group Robot increases by 7% (cf. figure 5). Difference is not statistically significant between the groups (Student’s t Test, t(0.7)=11.9, p = 0.4).

• Time to perform Peg Transfer: for Non-Resident participants, both groups decrease time to perform the Peg Transfer exercise after the learning session. Mean time to perform the exercise before the learning session is 3.1 ± 0.9 minutes, while mean time to perform the exercise after the learning session is, for group Classic of 2.6 ± 0.6 minutes and for group Robot of 1.9 ± 0.4 minutes (cf figure 6). Difference is significant between group Classic (M=2.6, SD=0.23) and group Robot (M=1.9, SD=0.18), Student t test, t(11.2)=2.09, p=0.05.

• Time to perform Peg Transfer: for Resident participants, both groups decrease time to perform the Peg Transfer exercise after the learning session. Mean time to perform the exercise before the learning session is of
3.3 ± 0.9 minutes, while mean time to perform the exercise after the learning session is, for group Residents Classic of 1.8 ± 0.6 minutes and for Residents Robot of 1.8 ± 0.4 minutes (cf figure 6).

4.2 Secondary Research Question: Does using a comanipulated interface for RALS improve performance compared with CLS?

- Time to perform training exercises: for Non-Resident participants, during the learning session, for exercise *Pea on a Peg*, group Classic starts for Trial n°1 with a mean time of 8.6 ± 1.9 minutes and ended for Trial n°5 with a mean time of 4 ± 2.4 minutes. Group Robot starts for Trial n°1 with a mean time of 4.4 ± 2.8 minutes and ends for Trial n°5 with a mean time of 3.1 ± 1.3 minutes. Difference of time taken across conditions is significantly different for trial n°1 at the p < 0.002 level [F(1,13) = 17.11, p = 0.001]. For exercise *Loops and Wire* no significant difference is observed between group Classic and group Robot. Group Classic starts for Trial n°1 with a mean time of 2.7 ± 0.7 minutes and ended for Trial n°5 with a mean time of 1.6 ± 1.4 minutes. Group Robot starts for Trial n°1 with a mean time of 3.2 ± 1.2 minutes and ends for Trial n°5 with a mean time of 1.5 ± 0.7 minutes (cf. figure 8). After verifying the normality and sphericity of data, and the independence of samples, a repeated measures analysis of variance was performed to test for effect of learning. This effect is significant for Group Classic for exercise *Pea on a Peg* (Anova, F(1,6) = 63.25, p = 0.0002), but not for group Robot (Anova, F(1,7) = 2.215, p = 0.18). For exercise *Loops and Wire*, the effect of learning is not significant for group Classic (Anova, F(1,6) = 3.68, p = 0.1), but it is for group Robot (Anova, F(2,5)=8.4, p = 0.025).

- Time to perform training exercises: for Resident participants, exercise *Pea on a peg* alone is performed. Group Residents Classic starts for Trial n°1 with a mean time of 4.7 ± 1 minutes and ends for Trial n°5 with a mean time of 2.9 ± 1.4 minutes. Group Robot starts for Trial n°1 with a mean time of 6.9 ± 2.8 minutes and ends for Trial n°5 with a mean time of 1.4 ± 1.6 minutes (cf. figure 8).

4.3 Exploratory Results: Self perceived comfort level with the task and observed performance

- Score for NASA TLX: for Non-Resident participants, they show an almost equal score for the two groups. Group Classic rates a mean score of 61 ± 6 and group Robot of 64 ± 12 (cf. figure 9).
• Score for NASA TLX: **for Resident participants**, both score very closely the task: 58 ± 18 for group Classic and 60 ± 19 for group Robot (cf. figure 9).

• Score at Peg Transfer exercise: **for Non-Resident participants**, mean score before the learning session for two groups is of 230 ± 81, after learning for group Classic of 163 ± 56, and after learning for group Robot of 158 ± 56. The two groups decrease their score in a similar way, showing equivalent improvement of skills.
Score at Peg Transfer exercise: for Resident participants, mean score before the learning session for the two groups is of $222 \pm 86$, after learning for group Classic of $111 \pm 59$, and after learning for group Robot of $147 \pm 38$. Group Classic decrease their score a little bit more than group Robot.

5 DISCUSSION

The measuring tools' and metrics’ results confirm their ability to show the differences in terms of skills and performance between the two experimental conditions (Co-RALS and CLS). They also confirm the absence of negative impact of training in Co-RALS on skills in CLS. Gaze-tracking enables to observe the subtle disparities in hand-eye coordination skills in CLS between training in Co-RALS and in CLS, although showing both train the human to perform laparoscopic surgery. Time recording showed a slightly significant difference in terms of the ability for speed in CLS for group trained in Co-RALS compared with group trained in CLS, but this result should be analyzed with care as before learning, group Robot showed better tim-wise skills than group Classic. The scores obtained in CLS by each group show no difference in achieved performance. Time recording when performing similar repeated exercises in Co-RALS and in CLS identifies that differences in the number of repetition needed before mastering the task, and the time performance that can be achieved for each condition depends on the type of task performed. The NASA TLX, finally, confirms that no supplementary workload was felt when performing in Co-RALS compared with CLS.

5.1 Main Research Question: Does learning laparoscopic surgery with a comanipulated robotic assistance result in development of skills in CLS?

5.1.1 Non-Resident participants. Both group Classic and group Robot decrease time taken to perform the Peg Transfer exercise. The difference between the groups is statistically significant, with group Robot taking on average 0.7 minutes less than group Classic after the learning session. Thus, group Robot seems to have developed better time-wise skills in laparoscopic surgery compared with group Classic. Gaze patterns between the two groups are very similar. Still, group Robot makes on average, after the learning session, slightly less and longer fixations in the three phases of the exercise: Grab, Transfer and Drop. These differences are not significant, but one can postulate that longer and repetitive learning sessions would expand them. Also, group Robot augmented, on average, the number of fixations on target before reaching it with the instrument in phases Transfer and Drop for targets n’1 to 5 while group Classic, on average,
slightly decreased it, showing little improvement for the first group, and consistency for the second. We interpret these results as a consequence of having to deal with the algorithm implemented in the robot, the viscous field forcing the participants in group Robot to remain active when performing the gesture. This may have increased their attention state, despite the fact that their score is almost equivalent to that of the group Classic. In other words, the results suggest that group Robot has developed skills in laparoscopic in a way that is not statistically different to those developed by group Classic in the same task. Without it being statistically significant, group Robot even has better results in terms of number of fixations on target, fixation rate per second and mean duration of fixations after learning than group Classic.

5.1.2 Resident Participants. Groups Classic and Robot, for Resident participants, took on average the same time to perform the Peg Transfer exercise after the learning session, suggesting they have equivalently improved their skills. Time taken to perform the Peg Transfer decreases equivalently for both group Classic and group Robot. The results from the small cohort of residents show a similar trend to that observed in the non-residents’ results. This suggests that the same results could be obtained on this experiment with larger groups of residents in medicine.

5.2 Secondary Research Question: Does using a comanipulated interface for RALS improve performance compared with CLS?

5.2.1 Non-Resident Participants. For Non-Resident participants, it appears that training with robotic assistance decreases the time taken to perform exercises that require precision and fine motor skills such as Pea on a Peg. The group of participants trained in Co-RALS took significantly less time to perform the first trial, and visibly less time to perform every other trial. Contrary to what was expected, for exercise Loops and Wire in which proprioceptive and instrument manipulating skills are required, comanipulated robot-assistance seems to hinder learning compared to no robotic assistance. The damping algorithm may, in this case, have disturbed the participants as it forces them to adapt to a new eye-hand coordination which has no advantage over the goal of this exercise. The implementation of an other algorithm such as virtual fixtures may help to improve performance for such an exercise but could have a negative impact on skill transfer to without robot mode. Further investigation is needed. Scores obtained at the NASA Task Load Index are very close for the two groups, not significantly different, showing an equivalent cognitive and physical load perceived when training either in CLS or in Co-RALS, suggesting it was as difficult to perform the learning session in CLS as it was in Co-RALS.

5.2.2 Resident Participants. For Resident participants, who had a better base level at laparoscopic surgery, more trials were required in RALS for Pea on a Peg exercise than for Non-Resident participants to attain the same level as in CLS, but better performance was attained with the same number of trials. This may be due to the fact that, as they were used to performing in CLS, they had to re-adapt to this new setting before they could achieve their best performance on it. No significant difference was observed regarding scores obtained at the NASA Task Load Index when training either in CLS or in Co-RALS. Same as for the main question of research, a similar trend than that of the novices is observed. This paves the way for more research on effect of Co-RALS on subjects with intermediate level in CLS, to assess whether their performance can be improved significantly compared with CLS.

5.3 Exploratory Results: Self perceived comfort level with the task and observed performance

5.3.1 Non-Resident Participants. Group Classic and Robot have on average equivalently improved their score at the Peg Transfer exercise before and after learning, with group Robot slightly better than group Classic. The mean score given
at the NASA TLX is almost equivalent for the two groups, showing no superior workload for group Robot compared with group Classic.

5.3.2 Resident Participants. Resident participants started with a score equivalent to the Non-Resident participants \(-222\) compared with \(230\) and ended with a slightly better score than the Non-Resident participants \(-111\) compared with \(153\) for group Classic and \(147\) compared with \(158\) for group Robot. No difference is observed in terms of perceived workload between group Robot and group Classic: the mean score at the NASA TLX is very close for the two groups.

6 CONCLUSION

The experiment conducted and its results, first, demonstrate that the measuring tools and metrics used and the exercises performed succeed in showing the learning process and performance of participants on tasks of robotic and classic laparoscopic surgery. The different metrics’ results also indicate the advantages and limits of Co-RALS for learning. A comanipulated interface for RALS seems to succeed in maintaining active learning of the motor skills of the user \(i.e.\) in the target task (CLS) while performing the robot-assisted task (Co-RALS), contrarily to what was observed by Blavier et al. with the da Vinci \[3\], \[4\]. However, these results confirm the motivations that lead to the development of comanipulation \[24\]: to support and assist the gesture to make it easier to perform without changing its characteristics.

Gaze-tracking and time recording permit observation of two different aspects of the learning process: the psycho-motor skills’ acquisition and the immediately visible performance. These two measuring tools taken together provide a comprehensive overview of what has been learned when training with robotic assistance. However, they go along with other measurements, whose purpose is to study the performance of the human-robot team. Thus, to evaluate the learning curve on exercises of laparoscopy either when training in Co-RALS or in CLS, we used time-recording also. We observed the learning curve of two exercises: one is learned faster and performed more rapidly in Co-RALS than in CLS, and the other is learned and performed more slowly with the robot compared to without. Hence, the superiority of the human-robot team’s performance in terms of the learning curve, seems to depend greatly on the type of exercise performed. Co-RALS enables to increase performance for exercises of fine motor skills and precision but for a task of proprioceptive and manipulating skills, Loops and Wire, it does not improve the learning process. The exploratory measures show other aspects of the human-robot interaction. The results obtained with the NASA TLX suggest that interacting with a robotic assistance to laparoscopic surgery results in an equivalent workload compared with performing CLS. The comparison of the peg transfer scores in CLS either after training in Co-RALS or in CLS demonstrate and equivalent performance between the two groups, confirming the previous results.

This study encourages to pursue research in human-robot interaction using quantitative metrics to qualify the conditions for interactions to be virtuous both for the human, and for the human-robot team’s performance. The results are also in favor of other interaction designs for RALS than the dominant one, Tele-RALS. Still, these findings present some limitations. First, the number of exercises performed are limited. Future research on comanipulated interfaces and RALS with longer, more complex and realistic tasks may show more clearly how it can benefit to students in surgery. Second, the number of Resident participants is small. A study with a greater number of Resident participants, with groups of different levels would show more precisely how the comanipulated robotic system impacts performance and develops psychomotor skills depending on the level in CLS. Also, our study only focuses on one aspect of laparoscopic surgery which is dexterity. One could imagine future research that would involve knowledge also: of anatomical structures, procedures, risks etc. Despite these limitations, our results pledge for more research on human-robot collaboration, which could lead to more adapted and hence more easily adopted technologies.
7 ACKNOWLEDGMENTS

The authors would like to thank Xavier Defieux and Sofiane Bendifallah, hospital practitioners who helped recruiting Resident participants and Gilles Bailly for his precious contributions to experimental design and article redaction. The authors would also like to thank all members of the Agathe team at Institut des Systèmes Intelligents et de Robotique (ISIR) who helped in editing the paper: especially Rémi Chalard, Ignacio Avellino, Angélina Bellicha, Etienne Moullet, Jesus Mago, Felix Péchereau, Fabien Vérité and Mégane Millan. The authors declare that they have no conflict of interest.

This work has been supported by the funds managed by the ANR within the Investissements d’Avenir program (Labex CAMI) under reference ANR-11-LABX-0004.

REFERENCES


