BEN-GURION UNIVERSITY OF THE NEGEV FACULTY OF ENGINEERING SCIENCES DEPARTMENT OF BIOMEDICAL ENGINEERING

# Surgeon-Centered Analysis of Robot-Assisted Needle Driving Under Different Force Feedback Conditions

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE M.Sc DEGREE

By: Lidor Bahar

February 2019

BEN-GURION UNIVERSITY OF THE NEGEV FACULTY OF ENGINEERING SCIENCES DEPARTMENT OF BIOMEDICAL ENGINEERING

# Surgeon-Centered Analysis of Robot-Assisted Needle Driving Under Different Force Feedback Conditions

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE **REQUIREMENTS FOR THE M.Sc DEGREE** 

By: Lidor Bahar Supervised by: Dr. Ilana Nisky

Author: .... 2 ....

Supervisor: .....

Chairman of Graduate Studies Committee:

Date: 17/05/2019

February 2019

Aug

## ABSTRACT

Robotic assisted minimally invasive surgery (RAMIS) systems provides many advantages to the surgeon and patient over open and standard laparoscopic surgery. However, haptic feedback, which is crucial for the success of many surgical procedures, is still an open challenge in RAMIS, and is missing in the majority of clinical RAMIS systems. Understanding the way that haptic feedback affects the surgeon's performance and learning can be useful in the development of haptic feedback algorithms and teleoperation control systems. In this study, we examined the performance of inexperienced participants under different haptic feedback conditions in a surgical needle driving via a soft homogeneous deformable object - an artificial tissue. We designed an experimental setup to characterize their movement trajectories and the forces that they applied on the artificial tissue. Participants first performed the task in an open condition, with a standard surgical needle holder, followed by teleoperation in one of three feedback conditions: (1) no haptic feedback, (2) haptic feedback based on position exchange, and (3) haptic feedback based on direct recording from a force sensor, and then again with the open needle holder. To quantify the effect of different force feedback conditions on the quality of needle driving, we developed novel metrics that assess the kinematics of needle driving and the tissue interaction forces. We combined our novel metrics with classical metrics such as task success and completion time. We analyze the final teleoperated performance in each condition, the improvement during teleoperation, and the aftereffect of teleoperation on the performance when using the open needle driver. We found that there is no significant difference in the final performance and in the aftereffect between the 3 conditions. Only the two conditions with force feedback presented statistically significant improvement during teleoperation in several of the metrics, but when we compared directly between the improvements in the three different feedback conditions none of the effects reached statistical significance. We discuss possible explanations for the relative similarity in performance. We conclude that we developed several new metrics for the quality of surgical needle driving, but even with these detailed metrics, the advantage of state of the art force feedback methods to tasks that require interaction with homogeneous soft tissue is questionable.

**Keywords:** teleoperation, force feedback, needle driving, robot assisted minimally invasive surgery

## Acknowledgments

I wish to thank my supervisor, Dr. Ilana Nisky for all the support, encouragement, and guidance during the past two years.

I would like to thank my friends and lab colleagues: Amit Milstein, Erez Sulimani, Reut Nomberg, Yarden Sharon, Chen Avraham, Danit Itzkovich and Mor Farajian for their friendship, moral support, and helpful advises.

#### Terminology

- DF direct feedback
- DOF degrees of freedom
- dVRK da-Vinci research kit
- F/T force / torque
- MIS minimally invasive surgery
- MTM master tool manipulator
- NF no feedback
- PE position exchange
- PSM patient side manipulator
- RAMIS robot assisted minimally invasive surgery

## Contents

1	INTRODUCTION			7	
2	MATERIALS AND METHODS			13	
	2.1	Needle	e driving task	13	
	2.2	Experiment setup		13	
		2.2.1	da-Vinci Research Kit - Hardware and Control	13	
		2.2.2	Open and Teleoperation Setup	17	
	2.3	Exper	imental Procedures	18	
	2.4	Data Analysis		20	
		2.4.1	Sampling and preprocessing	20	
		2.4.2	Segmentation	21	
		2.4.3	Metrics of performance	22	
		2.4.4	Statistical analysis	27	
3	RESULTS 2			29	
4	DISCUSSION		41		
<b>5</b>	CONCLUSION		48		
6	FUTURE WORK		49		
A	Appendix: Full results			60	
	A.1 Results for the full task			60	
	A.2 Results for the extraction subtask		65		
в	App	oendix	: Jerk metrics analysis	70	

## **1** INTRODUCTION

Robot Assisted Minimally Invasive Surgery (RAMIS) refers to minimally invasive surgical procedure aided by robots. RAMIS is the robotic development of laparoscopic surgery, in which the surgery is performed through several small incisions in the patient body. In RAMIS, the surgeons use a robotic manipulator to operate robotic instruments inside the patient's body. The surgical instrument (or end effector, e.g. gripper, scissors) follows the movement of the surgeons hands, usually with different motion scaling, filtering, and other possible manipulations or restrictions.

Minimally Invasive Surgery (MIS), yields faster recovery time for the patients, has aesthetical benefits, and for some cases have been shown to improve surgical outcome compared to open surgery[1]. RAMIS has all the advantages of MIS, with the additional key advantages such as 7 degrees-of-freedom (DOF), 3D high definition visual system, higher precision and accuracy, and more intuitive operation (e.g. preventing the Fulcrum effect). As a result, RAMIS has the potential to produce a better surgical outcome [1][2][3][4].

The field of general surgery has gone through a big change in the 1990s when several robotic teleoperation systems for minimally invasive surgery were presented [5]. At the moment, the most widely spread system in the market is the da-Vinci surgical system (Intuitive Surgical Inc.) with about 4400 installed systems worldwide [6] but other players are entering this market, such as the Senhance system (TransEnterix) that is approved for clinical use in Europe. In 2017, the number of surgical procedures using the widespread da-Vinci RAMIS system was ~877,000 [6], and this number has been consistently growing over the years [7][8][9].

Surgeons that were used to operate with their hands in full contact with the patient have moved to sit behind robotic interfaces that mediated between them and the patient. In this process, surgeons have lost the ability to use the crucial sense of touch, or haptic sense. The haptic sense is used in many surgical procedures for detecting lesions, orient inside the body and most importantly to facilitate the amount of force needed in each of the 7 DOF available in current teleoperation

systems [10][11].

The haptic sense is generally divided into two main modalities - kinesthetic and tactile [12]. Kinesthetic modality refers to forces and torques that are sensed in the muscles, tendons, and joints. This definition is then translated to haptic feedback, such that kinesthetic haptic feedback will be sensed by the kinesthetic modality. In this study, we will refer to kinesthetic haptic feedback as haptic feedback. In surgery, the haptic feedback is being used constantly by the medical staff, as the haptic sense has an important role in motor control and in many surgical procedures. Currently, the main commercial RAMIS system does not offer any haptic feedback to the surgeon, although the system is physically able to provide kinesthetic haptic feedback (Senhance system, TransAstrix; REVO-I Robotic Surgical System, Meere Company) [12] [13] [14].

As consequence of the lack of haptic feedback, surgeons have to estimate the forces based only on visual information (e.g. tissue deformation, color changes of tissue) and prior knowledge (e.g. the stiffness and deformability of the specific tissue). This estimation is worsened when the surgeon's field of view is obstructed. Prior literature stressed that estimating forces and torques using other modalities can cause an inaccurate estimation, and as a consequence surgeons can apply substantial forces on the tissue ending in unwanted injuries. This drawback is especially important when considering non-expert surgeons and trainees that are in the process of learning how to control and estimate the forces correctly, while the lack of feedback is making the use of the devices less intuitive [7][13][15].

The reasons for the lack of haptic feedback are mainly due to stability issues in closed loop teleoperation with force feedback and challenges in force estimation. Introducing haptic feedback in teleoperation can cause the system to get out of stability, which can lead to unwanted oscillations and difficulty of controlling the end effector [7] cause the system to get out of stability, studies on haptic interfaces and stability have managed to give haptic feedback and ensure a stable system, but at the expense of the transparency. Transparency is a measure of the fidelity of the teleoperation system. The transparency is defined as the ability to accurately display remote environment impedance to the operator [7] [16] [17]. Nisky et al. [17] suggested that transparency can be divided into 3 components, perceptual transparency, local motor transparency, and remote motor transparency. The second component, local motor transparency, is by definition achieved when the movement of the operator (position and force trajectories) does not change when the teleoperation system is replaced by an identity channel. Thus, the way the forces are rendered to the surgeon is another challenge, and understanding the effect of different conditions of force feedback on the surgeon (i.e. the movement of the operator) is a necessary step for optimizing force feedback in RAMIS.

A second reason for the lack of haptic feedback is the difficulty to estimate the forces that are applied at the end effector. Several haptic feedback algorithms were suggested to solve the challenge of force estimation. For example, position exchange is an algorithm that estimates the forces based on the error between the desired position of the robot and the actual (current) position of the robot [18]. Li et al [19] suggested sensorless force estimation using Gaussian processes regression (GPR) based on the motors' readouts and mathematical modeling of the Raven-II. Davland et al [20] designed an actuated force feedback-enabled laparoscopic instrument for robotic-assisted surgery. Anooshahpour et al [21] suggested a quasi-static model for the da-Vinci instrument and Rivero et al [22] suggested estimating the forces based on visual data (i.e. deformation of tissue). The suggestion of using visual information in force estimation is interesting since currently, this is the way that surgeons estimate the forces. Each one of the force feedback algorithms has a trade-off between system stability and transparency along with other limitation.

The effect of haptic feedback compared to condition with no haptic feedback is not yet fully understood [23][24][25][26]; in a meta-analysis on the effect of force feedback (i.e. kinesthetic feedback), Weber et al [27] found that force feedback reduced the amount of force applied, but in addition they found evidence that the effect of force feedback is smaller when depth perception is available (i.e. 3D vision). Another meta-analysis by Weber et al [28] showed that the effect of force feedback is task-depended; they found that when surgical tasks were given, the effect of force feedback was bigger, resulting in lower applied forces.

The results in studies that compared task performance between haptic feedback condition to no haptic feedback condition were inconclusive. Mahvash et al [29] compared the performance in palpation task using direct force feedback (measured by a force sensor located at the remote site), visual feedback of the forces on the screen, both direct and visual feedback, and without force feedback. They show mixed results in the question of whether force feedback yields better performance in two different perceptual palpation tasks. In a heart model, the force feedback condition with only haptic feedback yield better results compared to other conditions, but for a prostate model, they did not find differences between the conditions. In a follow-up experiment with similar conditions as Mahvash et al, Gwilliam et al [30] showed that the accuracy of experienced surgeons was higher when they received haptic feedback compared to visual force feedback or no feedback.

Santos-Carreras et al [31] compared participants performance in needle driving task in virtual reality for three different force feedback conditions, visual feedback (i.e. without haptic cues), 3 DOF force feedback and 6 DOF force feedback (forces and torques). It is important to note that the visual system in the experiment was a 2D screen (i.e. not 3D). Forces were rendered to the users by a set of equation modeling the forces in the task. In the experiment, participants performed the task using a surgical needle holder that was attached to a robotic interface. They tested three metrics - completion time, exit point error and maximum penetration depth. They show that there is a significant difference between force feedback conditions and visual feedback for exit point error and maximum penetration depth, but they did not find any benefits when adding torques to the force feedback.

In addition to metrics that evaluate task performance, we can define metrics to evaluate the operator hand kinematics and characterize operator's learning process. When manipulating the tool in a teleoperation system, we have different kinematics compared to free hand movements - even when looking on simple movements such as reaching to a target [32] or on a more complex movement such as needle driving [33]. In needle driving task, participants presented slower learning in teleoperation compared to needle driving preformed with needle holder as in open surgery. When learning to use a teleoperation system, participants presents learning curve over trials in path length, completion time and other kinematic and dynamic metrics. [32] [34].

Metrics that quantifies the human movement or performance are also used to differentiate between novices and experts surgeons [33] [35] or to evaluate the effect of training on performing surgical tasks [36]. Classical metrics, such as task completion time and path length were measured to asses the performance of participants during surgical tasks, the learning process was quantified using those metrics and learning curve were observed when looking at the performance trial by trial. [33] [37] Same metrics and other metrics are also used to assess the movement quality in rehabilitation process of patients. [38] Force and torque metrics were suggested to discriminate between novice and expert surgeons [39]. Sharon et al [40] showed that the change in instrument orientation during needle insertion movement, normalized by path length is different between expert surgeons and novices.

In order to asses the operator performance it is possible to define metrics that are based on research on the motor system. The motor system is complex and the different levels of movement execution are subjects of numerous studies [41] [42]. In order to execute a movement our motor system is engaged in a series of tasks that ends in the desired movement. There are infinite amount of ways the motor system can perform the same task, this is due to redundancy we have. As a result, modeling the motor system and its movements is a non-trivial task [43].

One prominent kinematic law of the human movement is the two-thirds power law for planar trajectories and its generalization to 3 dimensional movements, known as the one sixth power law. The two-thirds power law describes the relation between movement's curvature and speed, and the one-sixth power law add the dependency to torsion. Several studies have shown that the law applies for simple and complex movements [42] [44]. Sharon et al. [45] showed that surgeon expertise and the use of teleoperation effects the parameters of the one sixth power law.

Another suggested principle in human movement is the minimum jerk. Jerk is the 3rd derivative of the position, and studies showed that movement obey to the principle that human movement is optimize such that its jerk will be minimal [42]. Jerk is a measure to the smoothness of the movement, lower jerk is smother movement. Several metrics were suggested to quantify the smoothness of the human movement [46] [47][48][49].

In order to study how haptic feedback affects the motor system, and specifically how it affects the learning of the motor system, we need to quantify the performance as well as the different aspects of the movements using classic and novel metrics. We need to develop valid metrics for evaluation of the motor system performance. Using the metrics and trial by trial experiment with different kinesthetic force feedback conditions, we can asses the effect of the different haptic condition on human movement's learning and performance.

## 2 MATERIALS AND METHODS

#### 2.1 Needle driving task

We selected the needle driving surgical task for this experiment, and designed an experimental setup to characterize participants' movements and applied forces. Needle driving task encompasses a complex movement (movement in 6 DOF, divided into several subtasks) combined with interaction with soft tissue and it is a common procedure in surgeries. Tissue interaction forces can add important information to the operator when interacting with a soft object. The goal of needle driving is driving a surgical needle between two points through tissue using a surgical tool. The participants were requested to drive the needle while following the arc of the needle with the tooltip, maintaining tool tip velocity direction perpendicular to the needle curve. We divided the needle driving task into six subtasks (Fig. 1), (a) needle positioning, (b) needle insertion, (c) needle correction (optional), (d) repositioning, and (e) needle pulling. We refer to needle correction as needle insertion after repositioning the tool on the needle; therefore, this subtask is optional, and the participants were instructed to avoid corrections as much as possible.

#### 2.2 Experiment setup

#### 2.2.1 da-Vinci Research Kit - Hardware and Control

The experimental apparatus was the da-Vinci Research Kit (dVRK) [50]. The dVRK is based on first generation da-Vinci system hardware (Intuitive Surgical Inc.), and has open hardware controllers and open source code [51][50]. The dVRK teleoperation setup (Figure 2 A,B) included two Master Tool Manipulators (MTM), two Patient Side Manipulators (PSM), 4 controllers (one for each manipulator), foot pedals tray and high resolution stereo viewer. During teleoperation, the remote side manipulators (PSMs) follow the movement of the master side manipulators (MTMs) after scaling and proper changes to orientation. Each MTM is capable of providing



Figure 1: Needle driving task. We divided the needle driving task into 8 subtasks (A-H). On each trial, participants (A) grasped a needle and positioned it above the starting point, (B) transported the needle to the insertion point and positioned it, and (C) inserted the needle into the tissue (1st insertion). (D) If needed, participants repositioned the tool on the needle and continued in insertion (2nd insertion). Then they (F) repositioned the tool the needle for extraction, (F) grasped the needle for extraction, (G) extracted the needle, and (H) positioned the needle above the finish point.

6 DOF haptic feedback. In this experiment, we used 2 large needle drivers as patient side instruments. The 3D vision system consists of 2 HD cameras (FLIR Blackfly S cameras, BFS-U3-32S4C, 3.2MP; Edmund Optics Lenses, 16mm, f/1.8 Ci Series Fixed Focal Length Lens), and 2 HD flat screens (frame rate of 35 Hz, resolution of 1080 X 810). The dVRK controllers were connected to Ubuntu (UNIX) OS computer with an Intel Xeon E5-2630 v3 2.40GHz processor. The vision system was connected to Ubuntu OS computer with Intel Core i7-7700K 4.2GHz processor and NVIDIA Quadro P2000 5GB graphics card. The communication between the two computers was done over the Universitys local area network (LAN) and ROS multiple machines configuration. The maximal latency that was measured using this communication was 0.5 ms.



Figure 2: Experimental setup for teleoperation (A),(B) and for open needle driving (C). (A) Master side: the participant uses the two dVRK Master Tool Manipulators (MTM) to control the patient side tools, and receives 3D visual feedback via the 3D viewer. For position exchange (PE) and direct feedback (DF) conditions, the operator also receives 3 DOF haptic feedback. The foot pedal has to be pressed to enable teleoperation. (B) Patient side: needle driving is performed on the artificial tissue using the right tool of the patient-side manipulator (PSM). 2 HD cameras (not visible) are used to acquire the visual information that is presented to the user. 6 DOF pose of the tools is recorded by the robot. 10 needles are prepared for each block. Inset: a side view of the artificial tissue fixture. Forces and torques were recorded by the F/T sensor that is embedded in the fixture. (C) Open needle driving. The participant performed needle driving using a titanium needle holder. Magnetic transmitter records 6 DOF pose of the 2 magnetic sensors. The HD camera records the trial. Artificial tissue, fixture, and F/T sensor setup are identical to the teleoperated conditions.

The control mechanism of the dVRK teleoperation is depicted in figure 3. x and  $\dot{x}$  denotes end effector Cartesian space position and velocity respectively, and X denotes a vector of x and  $\dot{x}$ . q and  $\dot{q}$  denotes joints position and velocity respectively. M, P and U denotes MTM, PSM and User respectively. 'des' and 'cur' denotes desired and current. R, T, S and J denotes rotation matrix, transformation matrix, scaling and Jacobian respectively. Forward and inverse kinematics are denoted by Fwd and Inv respectively. f and  $\tau$  denotes Cartesian forces and motor torques respectively. The user moves the MTMs and receives 3DOF haptic feedback. The patient side (figure 3B) follows the master side (figure 3A) with proper scaling and transformation. Scaling  $S_{Tele} = 0.4$ , such that the user reaches the entire experimental workspace without consideration to workspace limits. The PSM control is depicted in figure 3D. The PSM interacts with the environment (i.e. needle, tissue).

When switch 1 is closed, the user receives position exchange (PE) haptic feedback. In PE haptic feedback, the PD controller (figure 3E) receives desired and current position and velocities, such that the error between them is used to calculate the force feedback. The  $K_P$  and  $K_D$  for the PD PE controller were selected, were tuned manually to roughly match the output forces to the forces recorded by the F/T sensor. The values are  $K_P = 1000 \frac{kg}{s^2}$  and  $K_D = 30 \frac{kg}{s}$ . When switch 2 is closed, the user receives direct feedback (DF) haptic feedback. The DF feedback is sampled by the force sensor in 500 Hz and is transmitted to the operator directly with a gain of  $G_{DF} = 0.7$ . We chose the 0.7 force scaling empirically as the largest gain in which participants were able to easily stabilize the system. Maximum latency in DF haptic feedback was  $\sim 20$  ms. We measured the latency experimentally by using left PSM to invoke the F/T sensor direct feedback to the right PSM. This way we could find the difference between the positions of both PSMs and measure the maximum latency. In both force feedback conditions, the forces were rendered to the user at 500 Hz. When both switches are open, the user did not receive feedback (NF). Both switches were never closed together. In addition, dVRK gravity compensation (GC) was enabled [50][51].



**Figure 3:** dVRK teleoperation control. (A) Master-side, (B) patient side, and (C) haptic feedback conditions: when switch 1 is closed, the user receives position exchange (PE) haptic feedback. When switch 2 is closed, the user receives direct feedback (DF) haptic feedback. When both switches are open, the user does not receive feedback (NF). In addition, dVRK gravity compensation (GC) is enabled. (D) Patient side dVRK controller (presented in B) (E) PD PE controller for the PE haptic feedback. Complete description is given in section 2.2.1

#### 2.2.2 Open and Teleoperation Setup

The experimental setup included two parts, (1) open needle driving, as in open surgery, and (2) teleoperated needle driving with the dVRK. The open needle driving setup (figure 2C) consisted of magnetic tracking sensors (TrakStar, Ascension Technologies, NDI), a titanium needle holder (Fine Needle holder, 16cm, Serrated, Titanium, World Precision Instruments), and a single HD camera (LifeCam Studio 1080p FHD WebCam, Microsoft). We instrumented the surgical needle holder with two magnetic 6DOF sensors. The sensors were mounted on the needle holder using custom made 3D printed fixtures that attached the sensors to the two rods of the needle holder. We calculated the tooltip position and orientation using the two magnetic sensors' position and orientation. The teleoperated needle driving setup (figure 2A,B) included the dVRK, 3D HD cameras and two large needle driver tools (Intuitive Surgical).

For both setups (open and teleoperated), we designed a 3D-printed custom fixture for the artificial tissue. The artificial tissue was made from a silicone molded homogenous piece (EcoFlex 00-30 mixed with 10% Silicone Thinner, Smooth-On inc.). The tissue was molded around 2 wooden sticks that were used to fix the tissue to the 3D printed fixture. The fixture and tissue were mounted on an acrylic plate that was fixed to a force/torque sensor (Nano 43 F/T sensor, ATI Industrial Automation) that measured tissue interaction forces and torques. The 3D-printed fixture could be rotated in 4 different right angles, to enable rotation of the tissue between the experimental blocks to allow for working on a fresh portion of the artificial tissue in each block. We used surgical needles for general surgery (GS-21, Covidien), 10 needles were placed on the fixture in each block for participants to use.

#### 2.3 Experimental Procedures

Thirty participants (N = 30) took part in the experiment after signing an informed consent form. The protocol and the form were approved by the Human Subject Research Committee of Ben Gurion University of the Negev, Beer-Sheva, Israel. All participants were right-handed, and they used the right needle driver to perform the needle driving, and the left needle driver (or their left hand in the open setup) to adjust the needle as needed.

Participants were asked to perform 120 trials of needle driving with the open and the teleoperated setups. The experimental protocol is depicted in figure 4. The protocol was divided into 12 blocks, and each block included 10 trials. After each block, the artificial tissue was rotated such that a fresh entrance and exit point were presented to the participant to prevent wearing of the silicone tissue. All the participants first performed 40 trials of open needle driving. We used the performance of each participant at the end of this baseline open part to compare with their performance at the subsequent parts of the experiment. In the second part of the experiment, all the participants performed 60 trials of needle driving task using the dVRK. Each group received one of three feedback conditions: (1) no haptic feedback (NF, N = 10), (2) direct sensing from a force sensor that was mounted under the tissue (DF, N = 10), and (3) position exchange based force feedback (PE, N = 10). After completing the teleoperation part, all the participants performed additional 20 trials using the open needle driving setup. This transition back to performing open needle driving after teleoperation represents a scenario that could happen in real life surgeries in case that the surgeon decides that the procedure cannot be safely completed with robotic assistance, and is used to assess the aftereffect of teleoperation with different feedback conditions on the performance in open needle driving.



**Figure 4:** Experimental protocol. Participants performed 120 trials of needle driving that were divided into 12 blocks. To avoid tissue fatigue, the entrance and exit point were refreshed in each block without changing the needle driving desired path by rotating the tissue fixture by 90 degrees in the horizontal plane. In our statistical analysis, we focused on 3 contrasts: learning, aftereffect and final performance. Learning: the effect of practice in teleoperation the difference in performance between early and late teleoperation. Aftereffect: the effect of teleoperation on open needle driving difference in performance at the end of teleoperation compared to the participants baseline the difference between the late teleoperation performance and the performance just before teleoperation.

Prior to the first part of open needle driving, and prior to the second part of teleoperation needle driving, participants viewed a video with instructions on how to perform the task correctly using the needle holder and using teleoperation respectively (as in [33]). Participants were asked to perform the needle insertion in one throw, and if needed, they could readjust their gripping of the needle and continue the insertion along the same curve of the needle. In addition, they were asked to work as quickly and as accurately as possible. Then, participants were asked to do a short practice. Before the open needle driving, participants practiced the use of the needle holder together with the surgical needle (without interaction with the tissue). Before the teleoperation part, participants were asked to adjust the system to their comfort (adjust chair and vision system's height and 3D vision of the environment was validated). Afterward, participants practiced several teleoperation exercises, to familiarize themselves with the dynamics of the robot. They were asked to: (1) move in a circle around the tissue and upwards, (2) reach in all directions - forwards, backward and to the sides, (3) touch the tissue and fixture using the teleoperation tool, (4) pick up the needle and release, and (5) pantomime the needle driving movements without touching the tissue.

After completing this practice, they began the test trials. After the first three trials of the open needle driving and first three trials in teleoperation needle driving part, participants received feedback on correct and incorrect performance; Emphasis was given on simultaneous wrist rotation during needle insertion and extraction, correct needle gripping, correct use of needle holder, and insert and exit of the needle at desired points. After each block, the tissue fixture was rotated, and participants received a short break.

#### 2.4 Data Analysis

#### 2.4.1 Sampling and preprocessing

In the open needle driving setup, we recorded the data in 120 Hz, and in the teleoperation setup we recorded in 500 Hz. We down-sampled and interpolated all the signals to 100 Hz. Both position and force data were interpolated using the shapepreserving piecewise cubic interpolation. Orientation data was interpolated using the spherical linear (slerp) quaternion interpolation method. We filtered tooltip position data with a 2nd order Butterworth low pass filter with cut-off frequency of  $F_c = 10Hz$  using Matlab *filtfilt()* function, resulting in zero phase filtering with a cut-off frequency of  $F_c = 8Hz$ . Velocity, acceleration, and jerk were calculated using numerical differentiation. After each differentiation, we filtered the signal using the same filter that was used for the tooltip signal.

#### 2.4.2 Segmentation

We defined trial start as the first interaction of the needle with the tissue, and the end of the trial as the final interaction with the tissue. We automatically segmented trial start and end based on force and tooltip position samples - the first and last of samples for which the force was higher than the noise threshold (th = 0.06[N]), and the tooltip position was inside the artificial tissue contour (i.e. there was interaction with the tissue, not with the fixture). We manually validated this trial segmentation based on force profiles and recorded trial video, and corrected erroneous segmentation due to accidental tissue interaction; Corrected segmentation samples were selected from a pool of samples that was marked by the automatic segmentation algorithm as possible segmentation samples.

We also automatically segmented each trial into trial subtasks (figure 1). When the gripper was closed and tissue interaction forces were above the noise threshold, the data point was classified as belonging to one of the needle driving segments (insertion, correction, or extraction). Insertion was defined as the first sequence of data points that belonged to the needle driving segments. Correction segments were classified if tooltip position was closer to needle entrance point rather than exit point. Extraction segments were defined as last sequence of data points belonging to the needle driving segment, or classified when tooltip position was closer to needle exit point rather than entrance point.

#### 2.4.3 Metrics of performance

We quantified participants performance using four classes of metrics: (I) task performance, (II) forces, (III) kinematics, and (IV) motor control grounded metrics. In the following section, we will review all the metrics that we used, including classical metrics and novel metrics that we developed in this study.

We used two metrics to quantify task performance. *Task completion time* was calculated as the cumulative time in which participants preformed the task:

$$Completion Time = t_{end} - t_{start}, (1)$$

where,  $t_{start}$  and  $t_{end}$  are the start and end time of the task respectively. *Exit point* error metric was measured as the Euclidean distance between desired exit point and actual exit point:

Exit Point Error = 
$$||\mathbf{x}_{desired} - \mathbf{x}_{actual}||,$$
 (2)

Where,  $\mathbf{x}_{desired}$  and  $\mathbf{x}_{actual}$  are vectors of the x, y data point in the plane of the tissue surface. The data of the actual and desired exit point was extracted using the images recorded by the cameras. Since the cameras and tissue in the teleoperation setup are fixed and identical between all trials, we could used the known dimension of the tissue to calculate the distance between them. Because the camera in the open needle driving setup was not fixed, we could not use the images to extract the exit point error, thus the exit point error metric was measured in the teleoperation needle driving setup only.

We used five metrics to quantify the forces that the participants applied on the tissue. We calculated the *total normalized force* as:

Total Normalized Force 
$$= \frac{1}{d_{ie}} \int_{t_1}^{t_2} |\mathbf{f}(\mathbf{t})| dt,$$
 (3)

where,  $|\mathbf{f}|$  denotes the total force,  $d_{ie}$  is the Euclidean distance between actual entrance and exit points. Since we used the Euclidean distance that was extracted only from the camera in the teleoperation setup, this metric is calculated only in the teleoperation setup. The metric quantifies the total forces that participants applied on the tissue as an indication of damage to the tissue. The rationale is that some amount of force is needed for the needle driving itself, but if there are substantial forces that were applied in incorrect directions the total cumulative forces will be higher. However, if the participants traveled a longer path inside the tissue due to inaccurate performance of the needle driving, larger forces will accumulate as well. Therefore, we normalized the cumulative forces by the distance traveled through the tissue.

Two additional metrics that quantified forces are the maximum force applied in the perpendicular component to the tissue (denoted as Z axis):

Max Force - Z axis = max 
$$(f_z)$$
, (4)

where,  $f_z$  denotes the force component in the vertical direction, and maximum force applied in the plane of the tissue (denoted as XY plane):

Max Force - XY plane = max (
$$|\mathbf{f}_{xy}|$$
), (5)

where,  $|\mathbf{f}_{xy}|$  denotes the force applied in the plane of the tissue. We quantified the maximum forces separately in the Z axis and XY plane separately because the forces in the XY plane are mainly designated for the needle insertion, while maximum force in Z component presents the damage to the tissue when pulling the tissue. The maximum torque around Z axis was calculated as:

Max Torque - Z axis = max 
$$(\tau_z)$$
, (6)

where,  $\tau_z$  denotes the torque around the vertical axis. This metric ideally should be zero, since the optimal needle driving movement is planar, and the needle should not rotate around the normal to the tissue direction.

The *force consistency* of consecutive movements was calculated as the sum of squared Euclidean cumulative distance between all force profiles to their mean profile:

Force Consistency = 
$$\frac{1}{N_t} \sum_{i=1}^{N_t} \sum_{n=0}^{N-1} (|\mathbf{f_i}(\mathbf{n}) - \mathbf{f_{mean}}(\mathbf{n})|_2)^2,$$
 (7)

where, **f** is the 3DOF tissue interaction forces, *i* denotes force profile index, and N is the number of samples within a single trajectory, and  $N_t = 5$  is the number of trajectories. Prior to the calculation of this metric we aligned the force trajectories using Dynamic Time Warping (DTW) to find the mean trajectory and to allow for calculating the distance between two trajectories.

We quantified movements kinematics using 5 metrics that calculates movement efficiency and consistency. We calculated the *path length* of the movement as:

Path Length = 
$$\sum_{i=1}^{N-1} |\mathbf{x}_{i+1} - \mathbf{x}_i|_2,$$
 (8)

Where, N is the number of samples. This is a classical metric [32] [34] [33] [35] [37], and lower values imply higher efficiency of the movement.

We calculated three metrics that quantify how close the participants were to performing the needle driving according to the instructions. For each trajectory, we fitted a circle that ideally should be with an identical diameter to that of the needle, and any deviation of the tip of the needle driver from that circle results in a movement that does not push the needle along its arc. We calculated the *Circle deviation* as the integrated distance of the projected path from the best fitted circle, normalized by the length of the projection of the path on the curve of the circle.

Circle Deviation = 
$$\frac{1}{\Theta_{arc}} \int_{\theta_1}^{\theta_2} \Delta s^2(t) d\theta$$
, (9)

where,  $\Delta s$  is the Euclidean distance between the fitted circle and the path projected on the fitted plane, and  $\Theta_{arc} = r_{needle} \Delta \theta$  is the total arc path in meters, where  $\Delta \theta$ is the difference between the angle of the tool tip at the beginning of the movement to the angle at the end of the movement on the arc of the fitted circle.

The *Plane deviation* quantifies the cumulative deviation from the movement plane the integrated distance of the trajectory from the plane that is fitted to the movement, normalized by the length of the path. We fitted the plane to the hand path using Matlab fitlm() function, we used the vertical axis (z axis) as the response (i.e. dependent variable), and x - y as the predictors (i.e. independent variables).

Plane Deviation 
$$= \frac{1}{PL} \int \Delta s^2(t) ds,$$
 (10)

where,  $\Delta s$  is the Euclidean distance between the fitted plane and the trajectory, and PL is the total path length in meters.



**Figure 5:** Illustration of the fitted plane and circle use in the kinematics metrics *circle deviation* and *plane deviation* for a single trial. In the figure we present the insertion subtask for one participants that received DF feedback, and the trial is recorded from the teleoperation setup. (A) The fitted plane and the path. (B) An example of a projected path (seen in 3 dimension in A) and the circle fitted to that path. (C) The distance of the path to the plane (i.e. shortest distance of each point). (D) The radius of the surgical needle (yellow) used in the experiment, together with the radius of the fitted circle (red) and the distance of the tooltip path from the center of the fitted plane.

The normalized angular path (eq. 11) quantifies technical aspects of needle driv-

ing, such as wrist rotation during needle driving [40].

Normalized Angular Path = 
$$\frac{1}{PL} \sum_{i=1}^{N-1} \Delta \theta_{i,i+1},$$
 (11)

where, PL is the path length,  $\Delta \theta_{i,i+1}$  is the angular rotation between two consecutive samples, and N is the number of samples. This metric was suggested in [40], where higher normalized angular path was associated with expertise and with learning of the needle driving with repeated performance of the task.

The trajectory consistency (eq. 7) of consecutive movements was calculated as the sum of squared Euclidean cumulative distances between all force profiles to their mean profile. We aligned force profiles using Dynamic Time Warping (DTW) (computed by Matlab dtw() function with the squared metric) to find the mean trajectory and to calculate the distance between each pair of trajectories.

Trajectory Consistency = 
$$\frac{1}{N_t} \sum_{i=1}^{N_t} \sum_{n=0}^{N_t} (|\mathbf{x}_i(\mathbf{n}) - \mathbf{x}_{mean}(\mathbf{n})|_2)^2,$$
 (12)

where, **x** is the 3 DOF position, *i* denotes trajectory index, and N is the number of samples within a single trajectory, and  $N_t = 5$  is the number of trajectories.

We used motor control grounded metrics to quantify to what extent the movements of the participants satisfy known laws in human motor control that were proposed for simpler movements. We focused on the speed-curvature-torsion power law [44][45] and the minimum jerk [46]. The speed-curvature-torsion power law is defined as:

$$v = \alpha \kappa^{\beta} |\tau|^{\gamma},\tag{13}$$

where v is the speed,  $\alpha$  is the velocity gain factor,  $\kappa$  is the curvature and  $\tau$  is the torsion.  $\beta$  and  $\gamma$  are the power constants, and in scribbling movements they were proposed to be equal to -1/3 and -1/6 respectively [44] (but may be different in needle-driving [45]. We used linear regression on the log-transformed data to fit the power law to the participants trajectories, and calculated the optimal  $\alpha$ ,  $\beta$ , and  $\gamma$  estimates.

In addition, we calculated the root mean squared jerk (eq. 14) of participants' movements.

$$\text{Jerk} = \sqrt{\frac{1}{t_2 - t_1}} \int_{t_1}^{t_2} \ddot{\mathbf{x}}(\mathbf{t})^2 dt, \qquad (14)$$

where,  $\ddot{\mathbf{x}}$  is the jerk of the movement. Jerk is a measure for movement smoothness, the lower the jerk, the smoother the movement. There are several suggested methods to calculate the jerk of a movement [46]. Following an analysis of the different methods on our data (see appendix B), we chose the method that normalizes the jerk by movement duration, and has units of jerk. As a result, this jerk metric is less correlated with completion time. Because this metric spans over several orders of magnitudes, we chose to analyze the logarithm of the metric.

#### 2.4.4 Statistical analysis

In our statistical analysis we compare three contrasts (figure 4): learning in teleoperation, aftereffect and final performance. In the learning contrast, we analyze the effect of practice in teleoperation, the difference in performance between early and late teleoperation. We define early and late as the median of the first 5 trials and median of last 5 trials in teleoperation, respectively. In aftereffect contrast, we analyze the effect of teleoperation on open needle driving, the difference in performance between just before (last 5 trials of 1st open needle driving) and just after teleoperation (first 5 trials of 2nd open needle driving). In final performance contrast, we analyze the performance at the end of teleoperation compared to participants baseline the difference between the late teleoperation (last 5 trials in teleoperation needle driving) performance and the performance just before teleoperation (last 5 trials in 1st open needle driving). We chose to use the medians of blocks of 5 trials in our statistical analysis to reduce the effect of trial-by-trial variability in the performance of the participants.

Since most of our metrics did not distribute normally. Therefore, we chose the median of each block as the statistic to quantify the metrics values. We used nonparametric statistical methods. For each one of the contrasts above, we used the two-sided Wilcoxon sign rank test to determine whether the difference between the two stages in each contrast was statistically different from zero. In addition, for each contrast separately, we used the Kruskal-Wallis test to determine whether we can reject the alternative hypothesis that the metrics describing the three force feedback conditions belong to the same distribution. Since none of the KW tests showed statistical significant differences no posthoc tests were needed.

### 3 RESULTS

To assess the performance of our participants in the needle driving task, we first visually examined the recorded trajectories of the tissue interaction forces and the kinematics of the patient side instrument tip. We compared the trajectories in the different dVRK teleoperation conditions, and the trajectories in the open needle driving with the surgical needle driver. Following the visual examination, we quantified the performance of the participants using novel and classical metrics. We grouped the metrics into the following categories: task performance, tissue interaction forces, kinematics, and motor control grounded metrics. Because most of the metrics were not normally distributed, we used non-parametric statistical tests to assess the effect of feedback conditions on each one of the metrics. Specifically, we focused on the effect of the different feedback conditions (no feedback NF, position exchange PE, and direct force feedback DF) on (i) the *learning* during teleoperation, (ii) the *final performance* in teleoperation, and (iii) the *aftereffect* of teleoperation on needle driving using an open surgical needle holder.

In figure 6, the tissue interaction forces and translation trajectories of a single participant in the first 5 trials (figure 6A,C) and the last 5 trials (figure 6B,D) are depicted. Visual examination of the force and translation trajectories suggests that the last 5 trials are more consistent (i.e. similar to each other) compared to the first 5 trials. Moreover, the trajectories seem to be less jerky and more planar (figure 6C,D). In Figure 7, Examples of tissue interaction forces as a function of time during a single trial in open needle driving (figures 7A,B) and teleoperation (figures 7C,D) are presented. When performing the task correctly (figures 7A,C), we identified a pattern in the tissue interaction forces. In the insertion part, the ideal driving trajectory is oriented along the x-y diagonal, and indeed the needle horizontal (x-y plane) forces are correlated. The vertical (z) component of the force vector is oriented in the negative vertical direction when the needle first penetrates the tissue, but the direction is reversed when the needle starts pointing towards the extraction point. During repositioning, there are occasional small forces due to transient interactions



**Figure 6:** An example of trajectories (lines) and tissue interaction forces (arrows) from the first 5 teleoperated trials (A) and last 5 trials (B) of participant from direct force (DF) feedback condition. The start of each trajectory is denoted by a diamond marker. Tissue interaction forces are represented by arrows, and each arrow represents the direction of the force at the specific sample together with the size of the force (length of the arrow). Different viewpoint of the same trajectories are presented in (C) and (D). The skin-toned flat cylinder represents the mock tissue specimen at scale. Red dotted line is presented to orient the viewer in space.

with the tissue. The extraction is a single fast movement of pulling the needle both in the horizontal plane and in the vertical axis, and as the movement progress, the vertical force component is more dominant compared to forces in horizontal plane. This pattern was consistent across participants in successful needle driving trials. For reference, examples of two trajectories of less successful trials are depicted in figure 7B for open needle driving and figure 7D for teleoperation. For both trajectories, participants did not succeed to insert the needle in one throw and needed a second throw to correct. The maximum tissue interaction force of the teleoperation example is higher compared to the correct example, and for the open needle driving the force trajectories have a different pattern and are less smooth.



**Figure 7:** Examples of force trajectories in each axis during needle-driving trials in open needle driving (A,B) and in teleoperation (C,D). We present force trajectory for a trial that was preformed correctly (open needle driving - A, teleoperation - C) and for a trial that was preformed incorrectly (open needle driving - B, teleoperation - D). In addition, the segments that were identified by the automatic segmentation algorithm are annotated.

In our analysis, we divided the needle driving task to several subtasks, and present here the results only for the insertion subtask. We focused on the *insertion* subtask because it is most important for the success of the overall needle driving, and its relative duration is highest (median relative duration and standard deviation of  $50.9\% \pm 10$  for insertion;  $16.6\% \pm 14$  for extraction;  $16\% \pm 14$  for repositioning) compared to the other subtasks. The full results of all subtasks are available in appendix A.

To compare participants performance between the three feedback conditions, we quantified three contrasts: (1) the learning during teleoperation (tele late - tele early), is presented in figure 9, (2) the aftereffect of each teleoperation condition on open needle driving (open 2 early - open 1 late) is presented in figure 11, and (3) the final performance of each participant in teleoperation compared to the participant's baseline performance (tele late - open 1 late) is presented in figure 10. For each metric, we present the median of the difference of each contrast and a non-parametric bootstrap 95% confidence interval for each of the metrics in each of the teleoperation conditions, along with markers denoting individual participants. Statistical results of all metrics for the three contrasts are presented in two tables: Wilcoxon sign rank test results are summarized in table 7 and Kruskal-Wallis test results are summarized in table 8.

First, we examine the *learning* during teleoperation contrast, in figure 8 we present a detailed look at the results of learning during teleoperation, in addition to the statistical analysis that is summarized in figure 9. Figure 8 presents the performances of the early stage teleoperation and late stage of teleoperation. For each stage, we present the median and non-parametric bootstrap 95% confidence interval for each of the metrics in each of the teleoperation conditions, along with markers denoting individual participants.

Looking at task performance metrics in teleoperation learning (panels A and B in figures 9 and 8), we observed significant improvement in completion time (effect of 2-5 seconds, table 7) for all conditions, without difference between the feedback conditions (figure 9B). There was no significant improvement in exit point error in teleoperation and no difference between the conditions (figure 9A). However, the radius of the exit point marker on the tissue was 2 mm, and thus, although we did not see a statistically significant difference between the groups, the force feedback

conditions (PE and DF) medians are closer to the marker radius compared to NF condition. Hence, it could be that participants did not try to improve their exit point accuracy more than the marker radius. Another important observation is that the dispersion of exit point error between participants was higher in the tele-early stage compared to tele-late stage (figure 8A), and this is more prominent in the force feedback conditions (DF and PE). So although there was no improvement in the groups' exit point error median, we think that participants with high exit error at the tele-early stage improved their exit point error accuracy, but we cannot provide quantitative statistical support to this observation.

Looking at the force metrics in teleoperation learning (panels C-G in figures 9) and 8), for the total normalized force metric, there was significant learning for all conditions without difference in learning between the three conditions (figure 9C). All participants reduced their applied forces during teleoperation, and therefore, in real surgery this would reduce the tissue damage. When isolating the maximum forces in the vertical component and horizontal plane, we did not see significant improvement in none of them (figure 9D,E), but when looking at the details (figure 8D,E), we can see that at the tele-early stage there was no visible difference between the conditions, and in the tele-late stage the dispersion for the force feedback conditions was reduced and the tendency was towards lowering the maximum forces, while the NF group did not reduce the median of the maximum forces. It is important to note that we did not find significant differences between the two stages in the conditions' median, nor in the interaction factor. In addition, when we looked at the maximum torque around the vertical axis (figure 9F), we saw that only PE feedback condition had significantly reduced its maximum torques, which ideally should be zero. We did not see a significant difference between the conditions, but we did see that the DF feedback condition had at first (Tele-early stage, figure 8F) higher maximum torques compared to tele-late stage (excluding 1 participant); this might be due to difficulty in stabilizing the system. Although participants did not receive feedback on the torques DOFs, most of them reduced the maximum torques around the vertical axis.



Figure 8: Comparison of the performance of the insertion subtask between early and late dVRK teleoperation. For each metric, median of first five (early) and last five (late) trials are presented. Each small marker represents an individual subject. The big marker with error bars shows the median of each group (NF, PE, and DF) and the 95% bootstrap confidence interval, respectively.

As for the consistency of tissue interaction forces, the *force consistency* metric, for both force feedback conditions, DF and PE, we saw a significantly improvement in the consistency of forces in learning during teleoperation (figure 9G). It is important to note that there was no statistically significant difference in learning between the conditions. Higher consistency of force profiles might suggest on a certain technique



**Figure 9:** Summary of the dVRK teleoperation *learning* during the insertion subtask. For each metric, median of difference between late teleoperation stage to early teleoperation stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.

that participants with force feedback acquired to help them perform the task.

Looking at the kinematics metrics in teleoperation learning (panels H-L in figures 9 and 8), we saw no significant difference between the feedback conditions. The movements of force feedback conditions (PE and DF) became significantly more consistent in the teleoperation learning process (figure 9L). The *trajectory consistency* for the DF group's median and variability was higher in the tele-early stage compared to the NF and PE conditions, while at the tele-late stage there was no visible difference between the groups (figure 8L); this might be because participants experienced difficulties in stabilizing the system in the DF condition at the beginning. There was a significant reduction during teleoperation learning in plane deviation for PE and NF (figure 9J), while there was no learning in DF. All 3 conditions presented lower dispersion in path length, deviation from circle, and deviation from plane in tele-late
stage compared to tele-early stage and there was no visible difference between the performances of the 3 conditions (figure 9H,I,J). We saw no learning for the angular path (figure 9K). Based on the results in [40] that showed that participants increased their angular path during the insertion task, we expected to see a similar increase in our study, but this was not the case.

Looking at the motor control oriented metrics in teleoperation learning (panels M-P in figures 9 and 8), participants' movements were generally consistent with the speed-curvature-torsion power law (mean  $R^2$  with standard deviation of the fit:  $R_{NF}^2 = 0.80 \pm 0.06$ ,  $R_{PE}^2 = 0.79 \pm 0.06$ ,  $R_{DF}^2 = 0.81 \pm 0.06$ ), and  $\beta$  and  $\gamma$  parameters had tendency towards their theoretical values (-1/3 and -1/6 respectively). Although there was no significant difference between the two stages, this result indicated that participants tendency was to perform movements that follow the 1/6 power law. The velocity gain factor,  $\alpha$ , showed that there is no visible difference between the 3 conditions in the two stages: it seems that participants did not move faster as the number of trials increases, as the tendency was not affected by the 3 conditions, nor was it reduced in the late stage. This result indicates that participants movement did not become smoother during teleoperation trials.

In the *final performance* contrast we quantified the performance at the late stage of teleoperation compared to the baseline performance at the late stage of the first open needle driving. The insertion completion time (figure 10A) of NF and DF conditions were significantly higher in teleoperation compared to open, while for PE condition there was no significant difference. As for the force metrics, there was no significant difference between the conditions for all force metrics. In the NF condition, participants applied significant more maximum force in the horizontal plane (effect size of 1.35 N) in tele-late stage compared to open 1-late (figure 10C), and there was no difference in maximum force for the force feedback conditions. For NF and PE conditions the consistency of forces (figure 10E) was lower in telelate stage compared to open-1 late stage, while for the DF condition, there was no



Figure 10: Summary of the dVRK teleoperation *final performance* compared to baseline for the insertion subtask. For each metric, median of difference between late teleoperation stage to late first open stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.

difference between the open needle driving and teleoperation needle driving.

In the final performance of the kinematic metrics participants path length (figure 10F), there was a significant difference between PE and DF conditions, meaning that PE force feedback group performed the task with lower path length compared to DF force feedback, when looking on the teleoperation needle driving compared to baseline. For the other kinematic metrics, there was no difference between the conditions in the final performance. Similarly, there was no significant difference between the feedback conditions in the motor control oriented metrics. Comparing the final performance in velocity gain factor  $\alpha$  metric (figure 10K), in NF and PE conditions participants movement was slower in teleoperation compared to open needle driving baseline. The jerk (figure 10N) for all conditions was significantly lower in the tele-



Figure 11: Summary of the open needle driving *aftereffect* for the insertion subtask. For each metric, median of difference between early stage in the second open needle driving to late stage in the first open needle driving are presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.

operation; this might be because the control system of the dVRK adds damping to participants movements.

For the *aftereffect* contrast (figure 11) we did not see aftereffect in open needle driving for none of the metrics. This means that the participants performance did not deteriorate due to the 60 trials of the task in teleoperation under any of the feedback conditions, and participants did not improve their performance in open needle driving after training the same task in teleoperation. The fact that the open performance did not deteriorate following teleoperation is encouraging as it suggests that if during surgical cases participants have to revert to open cases their performance would continue to be the same as before regardless to the force feedback condition.

	$ \begin{array}{ c c c c c c c c c c c c c c c c c c c$	F	Final performance							
Metric name		NF	PE	DF	NF	PE	DF	NF	PE	DF
	р	0.37	0.06	0.23	-	-	-	-	-	-
Exit point error	$\Delta$	0.50	0.95	1	-	-	-	-	Final performan         PE         -         0.43         -1.02         -         0.23         -0.37         0.131         -0.543         0.16         11.3         0.0097         0.023         0.0037         7.10 ⋅ 10 <sup>-6</sup> 0.037         7.2.64 ⋅ 10 <sup>-6</sup> 0.16         -7.89         0.62         -0.0010         0.0097         0.15         0.13         -0.021         0.16         -0.015         0.0019         1.4	-
Completion time	р	0.027	0.0019	0.0058	0.064	0.92	1	0.0058	inal performan PE - - 0.43 -1.02 - 0.23 -0.37 0.131 -0.543 0.16 11.3 0.013 -0.543 0.016 11.3 0.013 -364 0.0097 0.023 0.037 7.10 · 10 <sup>-6</sup> 0.037 2.64 · 10 <sup>-6</sup> 0.037 2.64 · 10 <sup>-6</sup> 0.16 -7.89 0.62 -0.0010 0.13 -0.62 -0.0010 0.015 0.13 -0.021 0.13 -0.021 0.13 -0.021 0.13 -0.021 0.13 -0.021 0.13 -0.021 0.015 0.13 -0.021 0.015 0.13 -0.021 0.015 0.015 0.016 -0.015 0.016 -0.015 0.017 0.017 0.016 -0.015 0.017 0.016 -0.015 0.017 0.017 0.016 -0.015 0.017 0.017 0.017 0.021 0.	0.037
Completion time	$\Delta$	1.99	3.8	5.47	-0.50	-0.079	-0.13	-1.62	-1.02	-1.64
Total Farma normalized	р	0.013	0.0039	0.0019	-	-	-	-	-	-
Total Force normalized	$\Delta$	0.46	0.66	0.98	-	-	-	Final perform           NF         PE           -         -           0.0058         0.43           -1.62         -1.02           -         -           0.0058         0.43           -1.62         -1.02           -         -           0.0064         0.23           -0.93         -0.37           0.00195         0.131           -1.35         -0.543           0.37         0.16           6.54         11.3           0.002         0.013           -968         -364           0.37         0.0097           0.011         0.023           0.37         0.0097           0.011         0.023           0.37         0.037           0.37         0.037           7         5.79 · 10 <sup>-6</sup> 7.10 · 10           0.32         0.62           -0.0042         -0.0010           0.32         0.62           -0.0042         -0.0010           0.013         0.0097           0.033         0.15           0.043         0.15           0.047	-	-
May fance 7 avia	р	0.69	0.16	0.16	0.69	0.16	1	0.064	0.23	0.37
Max Iorce - Z axis	$\Delta$	-0.20	0.31	0.41	-0.12	-0.23	0.10	-0.93	-0.37	0.84
Marsfanas XV alama	р	0.625	0.375	0.322	0.084	0.846	0.492	0.00195	0.131	0.695
Max force - XY plane	$\Delta$	-0.404	0.395	0.893	0.386	0.0344	0.35	-1.35	-0.543	-0.0909
Max torque - Z axis	р	0.69	0.0039	0.10	0.32	0.10	0.064	0.37	0.16	0.55
	$\Delta$	-3	7.9	5.2	3.03	11	6.87	6.54	11.3	9.36
DTW - force	р	0.19	0.027	0.037	0.84	0.43	0.10	0.002	0.013	0.19
	$\Delta$	451	1330	2140	16.6	-83.2	175	-968	-364	-192
Path length	р	0.084	0.10	0.019	0.23	0.037	1	0.37	0.0097	0.69
	$\Delta$	0.010	0.020	0.031	0.0082	0.010	0.0010	0.011	0.023	0.0020
Circle deviation	р	0.27	0.16	0.43	0.23	0.064	0.84	0.37	0.037	0.84
Circle deviation	$\Delta$	$4.78\cdot 10^{-6}$	$3.20\cdot 10^{-6}$	$1.20\cdot 10^{-6}$	$5.84\cdot10^{-6}$	$4.94\cdot 10^{-6}$	$8.82\cdot 10^{-7}$	$5.79\cdot 10^{-6}$	$7.10\cdot 10^{-6}$	$4.34\cdot 10^{-7}$
Plana doviation	р	0.027	0.013	0.55	0.84	0.19	0.55	0.43	0.037	0.69
I falle deviation	$\Delta$	$1.85\cdot10^{-6}$	$2.16\cdot 10^{-6}$	$1.75\cdot 10^{-6}$	$-4.81 \cdot 10^{-7}$	$1.82\cdot 10^{-6}$	$-8.20\cdot10^{-7}$	$8.62\cdot 10^{-7}7$	F         PE $F$ PE $  058$ $0.43$ $0$ $62$ $-1.02$ $  -$	$-1.80\cdot10^{-7}$
Angular path	р	0.10	0.55	0.27	0.69	0.92	0.13	0.77	0.16	0.16
Angular path	$\Delta$	-5.96	3.53	-9.44	-10.7	1.32	15.5	0.0058         -           -         -           0.0058         -           -         -           0.00195         -           -1.35         0.07           0.00195         -           -1.35         0.37           0.54         0.002           -968         0.37           0.011         0.37           5.79 · 10 <sup>-6</sup> 7.           0.43         8.62 · 10 <sup>-7</sup> 7           8.15         0.32           -0.0042         0.013           0.083         0.77           -0.0047         0.49           -0.0058         0.0019           1.22         1.22	-7.89	16.6
DTW position	р	0.10	0.019	0.027	0.77	0.43	0.92	0.32	0.62	0.10
DIW - position	$\Delta$	0.010	0.020	0.026	0.0021	$-4.28\cdot10^{-5}$	0.0014	-0.0042	-0.0010	-0.0041
1/6 power law o	р	0.32	0.084	0.13	0.77	0.32	0.23	0.013	0.0097	0.084
1/0 power raw a	$\Delta$	0.035	0.038	0.059	0.016	0.038	0.049	0.083	0.15	0.061
1/6 power law B	р	0.32	0.013	0.32	0.92	0.77	0.37	0.77	0.13	0.13
	$\Delta$	-0.0057	-0.019	-0.011	0.0040	0.0048	-0.0087	-0.0047	-0.021	-0.012
1/6 power law o	р	0.13	0.13	0.019	1	0.69	0.84	0.49	0.16	0.84
1/0 power raw 'y	$\Delta$	-0.015	-0.0095	-0.022	-0.0040	-0.00061	-0.0016	NF         PE           -         -           0.0058         0.43           -1.62         -1.02           -         -           0.0064         0.23           -0.93         -0.37           0.00195         0.131           -1.35         -0.543           0.37         0.16           6.54         11.3           0.002         0.013           -968         -364           0.37         0.0097           0.011         0.023           0.37         0.0097           0.011         0.023           0.37         0.0097           0.011         0.023           0.37         0.037           5.79 · 10 <sup>-6</sup> 7.10 · 10           0.43         0.037           8.62 · 10 <sup>-77</sup> 2.64 · 10           0.77         0.16           8.15         -7.89           0.32         0.62           -0.0042         -0.0016           0.133         -0.0047           0.034         0.013           0.035         0.15           0.77         0.13           -0.0047	-0.015	0.0041
Iork	$\begin{array}{c c} & & \\ & p \\ & \Delta \\ & p \\ & p \\ & \Delta \\ & p \\ & \rho \\ & \Delta \\ & p \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \rho \\ & \rho \\ & \rho \\ & \Delta \\ & \rho \\ & \Delta \\ & \rho \\ & \rho$	0.69	0.37	0.27	0.16	0.19	0.37	0.0019	0.0019	0.0019
JCIK	$\Delta$	-0.071	-0.070	-0.14	0.11	0.21	0.099	1.22	1.4	0.86

 Table 1: Statistical analysis - Wilcoxon sign rank

Matuia mana		Learning		Aftereffect			Final performance			
Metric name		NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF
	p model and $\chi^2$	p =	0.80, $\chi^2 =$	0.43		-			-	
Exit point error	р	0.84	0.82	1	-	-	-	-	-	-
	$\Delta$	-2.2	-2.3	-0.1	-	-	-	-	-	-
	p model and $\chi^2$	p =	0.17, $\chi^2 =$	3.54	p =	0.62, $\chi^2 =$	0.94	p =0	1) performant NF-DF - - - - - - - - - - - - -	1.04
Completion time	р	0.93	0.17	0.32	0.64	0.71	0.99	0.69	0.99	0.61
	$\Delta$	-1.4	-7	-5.6	-3.5	-3.1	0.4	-3.2	l performa NF-DF - - - - - - - - - - - - -	3.7
	p model and $\chi^2$	p =0	$0.093, \chi^2 =$	4.75		-			-	
Total Force normalized	р	0.77	0.083	0.31	-	-	-	-	-	-
	$\Delta$	-2.7	-8.4	-5.7	-	-	-	-	-	-
	p model and $\chi^2$	p =	$0.34, \chi^2 =$	2.14	p =	0.44, $\chi^2 =$	1.64	p =	0.082, $\chi^2$	= 5
Max force - Z axis	р	0.47	0.37	0.98	0.42	0.93	0.64	0.50	0.065	0.50
	$\Delta$	-4.6	-5.3	-0.7	4.9	1.4	-3.5	Final perform NF-PE NF-DF - - - - - - - - - - - - -	-4.4	
	p model and $\chi^2$	p =0	$.358, \chi^2 =$	2.05	p =0	.607, $\chi^2 =$	0.999	p =0	Final performa PE NF-DF 	= 4.22
Max force - XY plane	р	0.663	0.329	0.842	0.583	0.923	0.815	0.257	0.13	0.933
	$\Delta$	-3.4	-5.6	-2.2	3.9	1.5	-2.4	-6.2	inal performa NF-DF - - - - - - - - - - - - -	-1.4
	p model and $\chi^2$	p =	= 0.22, $\chi^2$	= 3	p =	0.91, $\chi^2 =$	0.17	p =	$0.92, \chi^2 =$	0.15
Max torque - Z axis	р	0.24	0.35	0.97	0.91	0.99	0.95	0.93	1	0.94
	$\Delta$	-6.3	-5.4	0.9	1.6	0.5	-1.1	-1.4	-0.1	1.3
	p model and $\chi^2$	p =	$0.34, \chi^2 =$	2.11	p =	0.27, $\chi^2 =$	2.60	p =	$0.14, \chi^2 =$	3.86
DTW - force	р	0.55	0.34	0.93	0.88	0.51	0.25	0.35	0.13	0.85
	$\Delta$	-4.1	-5.5	-1.4	1.9	-4.3	-6.2	-5.4	-7.5	-2.1
	p model and $\chi^2$	p = 0	$0.052, \chi^2 =$	5.91	p =	0.41, $\chi^2 =$	1.77	$\mathbf{p} = 0$	$0.028, \chi^2$	= 7.1
Path length	р	0.97	0.11	0.069	0.92	0.63	0.39	0.23	0.56	0.022
	$\Delta$	0.9	-7.8	-8.7	-1.5	3.6	5.1	-6.4	Final perform NF-PE NF-DF 	10.4
	p model and $\chi^2$	p = 0	0.87, $\chi^2 =$	0.266	p =	0.20, $\chi^2 =$	3.19	p =	$0.11, \chi^2 =$	4.41
Circle deviation	р	0.98	0.86	0.94	0.74	0.55	0.17	0.41	0.69	0.093
	$\Delta$	0.7	2	1.3	-2.9	4.1	7	-5	al performa           NF-DF           -           0.59, $\chi^2$ =           0.99           0.5           -           -           0.121, $\chi^2$ =           0.121, $\chi^2$ =           0.121, $\chi^2$ =           0.13           -7.6           0.92, $\chi^2$ =           1           0.14, $\chi^2$ =           0.56           4           0.11, $\chi^2$ =           0.56           4           0.10, $\chi^2$ =           0.51           4.3           0.11, $\chi^2$ =           0.72           -3           0.46, $\chi^2$ =           0.99           0.2           0.67, $\chi^2$ =           0.99           0.2           0.72           -3           0.72           -3           0.72           -3           0.72           -3           0.72           -3           0.72           -3           0.66, $\chi^2$ =           0.72	8.2
	p model and $\chi^2$	p =	0.44, $\chi^2 =$	1.61	p =	0.21, $\chi^2 =$	3.07	p =	0.10, $\chi^2 =$	4.45
Plane deviation	р	0.80	0.41	0.80	0.63	0.67	0.18	0.56	0.51	0.088
	Δ	2.5	5	2.5	-3.6	3.3	6.9	-4	Final perform PE NF-DF - - - - - - - - - - - - -	8.3
	p model and $\chi^2$	p =	0.26, $\chi^2 =$	2.66	p =	0.32, $\chi^2 =$	2.25	Final performal NF-PE NF-DF 	4.33	
Angular path	р	0.31	0.99	0.35	0.99	0.42	0.37	0.39	0.72	0.098
	$\Delta$	-5.7	-0.3	5.4	0.4	-4.9	-5.3	5.1	-3	-8.1
	p model and $\chi^2$	p =	0.38, $\chi^2 =$	1.91	p =	0.48, $\chi^2 =$	1.46	p =	0.46, $\chi^2 =$	1.52
DTW - position	р	0.85	0.35	0.67	0.63	0.97	0.48	0.55	0.99	0.51
	$\Delta$	-2.1	-5.4	-3.3	3.6	-0.9	-4.5	-4.1	0.2	4.3
	p model and $\chi^2$	p =	0.88, $\chi^2 =$	0.23	p =	0.68, $\chi^2 =$	0.75	p =	0.67, $\chi^2 =$	0.77
1/6 power law $\alpha$	р	0.95	0.88	0.97	0.67	0.81	0.97	0.74	0.99	0.71
	Δ	-1.1	-1.9	-0.8	-3.3	-2.4	0.9	-2.9	0.2	3.1
	p model and $\chi^2$	p =	0.46, $\chi^2 =$	1.53	p =	0.66, $\chi^2 =$	0.80	p =	0.51, $\chi^2 =$	1.32
1/6 power law $\beta$	р	0.50	0.99	0.56	0.84	0.64	0.94	0.48	0.77	0.89
	Δ	4.4	0.4	-4	2.2	3.5	1.3	4.5	2.7	-1.8
	p model and $\chi^2$	p =	$0.56, \chi^2 =$	1.14	p =	$0.98, \chi^2 =$	0.023	p =	$0.31, \chi^2 =$	2.32
1/6 power law $\gamma$	р	0.85	0.85	0.53	0.98	0.99	0.99	0.72	0.72	0.28
	Δ	-2.1	2.1	4.2	-0.6	-0.3	0.3	3	Final perform.           NF-PE         NF-DF           -         -           -         -           -         -           0.69         0.99           -3.2         0.5           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           -         -           0.50         0.032           -5.4         -7.5           -5.5         0.13           -5.4         -7.5           -5         3.2           -5         3.2           -5         3.2           -5         3.2           -5         3.2           -5         3.2           -7         -5	-6
	p model and $\chi^2$	p = 0	0.88, $\chi^2 =$	0.235	p =	0.37, $\chi^2 =$	1.94	$\begin{array}{c c} \mathrm{NF-PE} & \mathrm{NF-DF} \\ & & - \\$	= 5.61	
Jerk	р	1	0.90	0.91	0.55	0.95	0.38	0.81	0.21	0.057
	Δ	0.1	1.7	1.6	-4.1	1.1	5.2	-2.4	6.6	9

 Table 2: Statistical analysis - Kruskal-Wallis test

## 4 DISCUSSION

We examined the effect of feedback conditions on several aspects of task performance and movement. We used a task of surgical needle driving to evaluate this effect on participants in two setups, (1) teleoperation as in RAMIS, and (2) using a needle holder as in open surgery. We designed an experimental setup, such that participants performed the task in the two setups while we simultaneously recorded their position, orientation, and tissue interaction forces. Force feedback conditions include two simple force feedback approaches (1) position exchange (PE) force feedback, and (2) direct force (DF) feedback sensed from an F/T force sensor. The third condition was no force (NF) feedback condition. We designed a protocol to test the learning during teleoperation, the final performance in teleoperation and the aftereffect of teleoperation on open needle driving.

The experiment apparatus and protocol we designed enabled us to query different aspect of human movement combined with forces for complex movement in a surgical task in RAMIS. We developed new metrics to assess the quality of surgical needle driving: trajectory consistency, force consistency, circle deviation, plane deviation. We also used recently proposed as well as classical metrics to assess performance: exit point error, completion time, total normalized force, max force in z and x - y, max torque z axis, path length, angular path, speed-curvature-torsion power law, and jerk.

In a study with 30 participants, we found that in several of the metrics, such as force consistency and maximum applied torques in the vertical direction, the two conditions with force feedback presented statistically significant improvement during teleoperation (i.e. learning). In addition, in some of the metrics, such as task completion time and total normalized force, there was an improvement in all the force feedback conditions. However, when we compared directly between the learning in the three different feedback conditions, none of the effects reached statistical significance. Moreover, we did not find any statistically significant differences between the 3 force feedback conditions in the final performance in teleoperation (except one difference between DF and PE in path length) and the aftereffect. Therefore, we conclude that in our detailed analysis of the performance of needle driving through soft tissue we did not find a benefit to presenting novice participants with force feedback. Hence, our data suggests that the advantage of state of the art force feedback methods to tasks that require interaction with homogeneous soft tissue is questionable.

Up to performing this research, several studies compared conditions with force feedback (usually one algorithm of force feedback) to no force feedback condition, and usually added another condition with the force feedback as visual information [30][52][53][31][25]. Results were inconclusive and were task dependent. We expected to find statistically significant differences between the three force feedback conditions, and more specifically, we expected to find worse performance at the end of teleoperation without force feedback compared to the other two force feedback conditions, as in [25][54][30], and applied forces in [23], [31]. Surprisingly, we did not see such differences for all the contrasts and metrics except one, consistently with task completion time in [55][31], and task error in [23].

One possible explanation is that the 3D high-definition visual feedback compensated for the missing in the no force feedback (NF) and the poor haptic feedback in the position exchange force feedback (PE) conditions. The 3D visual compensation is more prominent when interacting with a deformable object such as artificial silicone tissue. Indeed, previous studies showed the importance of 3D vision [27]. Another explanation might be that after 60 trials in the novel environments in our experiment the human motor system adjusted to the new conditions and perform the same. In [33], learning curve of task completion time of non-medical novice participants in needle driving task did not seem to reach a plateau after 80 trials, while for path length it seems that the majority of learning improvement was in the first 20 trials. Another possibility is that 60 trials in teleoperation were not sufficient to the learning of the motor system, and there is potential for further improvement in task performance that would eventually reveal advantages to one of the force feedback conditions. This explanation is consistent with [32], where learning was observed even after 200 trials in a simple reaching task in teleoperation.

Another possible explanation for our failure to find statistically significant differences between the force feedback conditions might be the high between participants variability. This variability could be a consequence of the location of needle grasping, or variability in entrance and exit points. Large variability in the sample reduces the power of statistical tests. This indicates that larger sample sizes or adding constraints on the task may be useful to increase power in future studies. To partially mitigate the influence of between-participants variability, for each contrast we compared the performance of the participants to their relevant baseline performance. One more reason for the large variability is the human movement variability, which allows to successfully perform a given task with different trajectories. Generally, variability is unavoidable in the motor system, and is even considered a virtue of the sensorimotor system [56] [57] [58] [59] [60] [61] [61] [62].

With repetitions of the task, the decrease in trajectory consistency and force consistency metrics showed that the within-participant variability in consecutive trials was lower at the late stage of the experiment. This suggests that unlike in simpler movements [63] [64], participants gradually progressed towards adopting a stereotypic trajectory to perform the task.

In our experiment, we implemented two force feedback conditions, (1) position exchange force feedback and (2) direct sensing from a force sensor that was attached to the tissue plate. In addition, all participants in all 3 conditions received gravity compensation from the current available open source dVRK code. It should be noted that the current implementation of gravity compensation is simplest [51] and it could be further improved by more advanced algorithms [65]. It is possible that with better gravity compensation the participants could perform better and this could have allowed for observing more pronounced differences between the difference force feedback conditions.

In position exchange force feedback, we estimated the applied forces based on the

error between desired and current state of the robot tooltip [18]. While we did not perform an analytic analysis of stability in our force feedback implementations, we could easily see that the position exchange force feedback remained stable throughout the experiment. However, the force feedback did not feel real, as the forces when interacting with the tissue felt lower compared to DF condition and open needle driving. The position exchange form of force feedback is better when interacting with a stiff object rather than a soft object. Since the interaction in our experiment is with a deformable tissue, the forces felt less realistic and smaller than the actual applied forces. In addition, because there was no dynamic compensation, participants felt forces when moving the tool freely without any interaction with an object. This two issues likely resulted in a poor quality of force feedback in the PE condition.

In the direct force feedback condition, participants received the forces directly from a force sensor located under the tissue. The force sensor recorded the tissue interaction forces with high accuracy, and as a result, the forces felt more realistic than in the PE condition. However, occasionally, we observed the tool starting to get out of stability. For most cases, participants succeed to stabilize the system, and if they did not succeed they stopped teleoperation for a brief moment and returned to the task. To mitigate this problem, we empirically tuned a gain of 0.7 for the direct force feedback. However, in some cases, participants were still not able to stabilize the system. In general, although the direct feedback condition had stability issues, participants succeed to stabilize the direct force feedback condition. In some metrics, the DF condition median was visibly difference compared to the PE and NF condition at the early stage trials in teleoperation (e.g. completion time, force consistency and trajectory consistency), while the difference was not visible in the late stage.

In the *final performance* contrast, our aim was to compare the performances between the three conditions at the late stage in teleoperation. We chose to normalize each participants performances to his or her baseline performances at the late stage of open needle driving. As a result, in this contrast, we compared between the performances in open needle driving to the performances in teleoperation. Comparison between the two setups, in which the participants perform the task with different tools, is not trivial. However, because we compare the way the human operator manipulates the tool end effector (i.e. needle holder tooltip and dVRK large needle driver tooltip), and the metrics we used are relevant for the end effector we believe that this kind of comparison is valid. We did not see statistically significant difference between the feedback conditions in *final performance* (except one difference in path length metric), but we did see that for all metrics the jerk in the teleoperation was lower compared to open needle driving and that forces were for most cases similar in teleoperation late stage and open needle driving.

# 5 CONCLUSION

In the current study, we failed to find statistically significant differences between the feedback conditions in needle driving through soft tissue. Both our implementations of simple force feedback algorithms did not significantly improve the kinematics and tissue interaction forces of the human operator compared to no force feedback, and there was no difference between the two force feedback algorithms. This may be a result of the poor quality of our force feedback algorithms for soft tissue interaction, or due to sufficient force cues available from visual information about tissue deformation.

Detailed understanding of the effect of haptic feedback on the operators movement and tissue interaction forces is important for developing future guidelines for human-centered design of teleoperated RAMIS systems with force feedback that will eventually provide advantage over unilateral teleoperation. We think that better, human-centered, force feedback algorithms have the potential to improve performance and movement quality compared to no feedback at all or algorithms that don't consider the human operator as part of the control loop.

# 6 FUTURE WORK

In the future, it will be interesting to perform a similar experiment with experienced and novice surgeons that are familiar with surgery and RAMIS and to see if one of the force feedback condition will be beneficial for them. In this study, we implemented two simple force feedback algorithms, it is important to investigate the performance of participants with more force feedback algorithms [31][66][67][21][22] [68], and more importantly, in a condition with human-centered force feedback algorithm. Another aspect to study is to see if one of the explanations that we suggested above, for the lack of differences between the conditions, affected the results of this study. It is possible to perform similar experiment with different vision conditions, and particularly the 3D vision compared to 2D vision; To investigate the learning curves and perform the research with more trials; to add more constraints to the needle driving complex task; or to give feedback on performances in real time that will encourage the participants to perform better in the task.

## REFERENCES

- A. Szold, R. Bergamaschi, I. Broeders, J. Dankelman, A. Forgione, T. Lang, A. Melzer, Y. Mintz, S. Morales-Conde, M. Rhodes, R. Satava, C.-N. Tang, R. Vilallonga, and =European Association of Endoscopic Surgeons, "European Association of Endoscopic Surgeons (EAES) consensus statement on the use of robotics in general surgery," *Surgical endoscopy*, vol. 29, no. 2, pp. 253–288, Feb. 2015. [Online]. Available: http://europepmc.org/abstract/MED/25380708
- [2] M. E. Hagen, J. J. Meehan, I. Inan, and P. Morel, "Visual clues act as a substitute for haptic feedback in robotic surgery," *Surgical Endoscopy*, vol. 22, no. 6, pp. 1505–1508, Jun. 2008. [Online]. Available: https://doi.org/10.1007/s00464-007-9683-0
- [3] A. Garcia-Ruiz, "Manual vs Robotically Assisted Laparoscopic Surgery in the Performance of Basic Manipulation and Suturing Tasks," Archives of Surgery ( Formerly : Ama Archives of Surgery ) \*, vol. 133, no. 9, p. 957, 1998.
- [4] S. Maeso, M. Reza, J. A. Mayol, J. A. Blasco, and M. Guerra, "Efficacy of the Da Vinci Surgical System in Abdominal Surgery Compared With That of Laparoscopy," Annals of Surgery \*, vol. 252, no. 2, pp. 254–262, 2010.
- [5] J. H. Palep, "Robotic assisted minimally invasive surgery," Journal of Minimal Access Surgery, vol. 5, no. 1, pp. 1–7, 2009. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2699074/
- [6] I. S. Inc., "Annual report 2017," Intuitive Surgical Inc., Sunnyvale (CA), Tech. Rep., 2017.
- [7] N. Enayati, E. De Momi, and G. Ferrigno, "Haptics in Robot-Assisted Surgery: Challenges and Benefits," *IEEE Reviews in Biomedical Engineering*, vol. 9, pp. 49–65, 2016.

- [8] C. Freschi, V. Ferrari, F. Melfi, M. Ferrari, F. Mosca, and A. Cuschieri, "Technical review of the da Vinci surgical telemanipulator," *The International Journal* of Medical Robotics and Computer Assisted Surgery, vol. 9, no. 4, pp. 396–406, 2013.
- B. S. Peters, P. R. Armijo, C. Krause, S. A. Choudhury, and D. Oleynikov, "Review of emerging surgical robotic technology," *Surgical Endoscopy*, vol. 32, no. 4, pp. 1636–1655, Apr. 2018. [Online]. Available: https://doi.org/10.1007/s00464-018-6079-2
- [10] A. Trejos, J. Jayender, M. Perri, M. Naish, R. Patel, and R. Malthaner, "Robot-assisted Tactile Sensing for Minimally Invasive Tumor Localization," *The International Journal of Robotics Research*, vol. 28, no. 9, pp. 1118–1133, May 2009. [Online]. Available: https://doi.org/10.1177/0278364909101136
- [11] S. De, J. Rosen, A. Dagan, B. Hannaford, P. Swanson, and M. Sinanan, "Assessment of Tissue Damage due to Mechanical Stresses," *The International Journal of Robotics Research*, vol. 26, no. 11-12, pp. 1159–1171, Nov. 2007.
  [Online]. Available: https://doi.org/10.1177/0278364907082847
- [12] H. Culbertson, S. B. Schorr, and A. M. Okamura, "Haptics: The Present and Future of Artificial Touch Sensation," Annual Review of Control, Robotics, and Autonomous Systems, 2018. [Online]. Available: https://doi.org/10.1146/annurev-control-060117-105043
- [13] A. M. Okamura, "Haptic feedback in robot-assisted minimally invasive surgery," *Current opinion in urology*, vol. 19, no. 1, pp. 102–107, Jan. 2009.
- P. P. Rao, "Robotic surgery: new robots and finally some real competition!" World Journal of Urology, vol. 36, no. 4, pp. 537–541, Apr. 2018. [Online]. Available: https://doi.org/10.1007/s00345-018-2213-y

- [15] P. J. Johnson, D. E. Schmidt, and U. Duvvuri, "Output control of da Vinci surgical system's surgical graspers," *Journal of Surgical Research*, vol. 186, no. 1, pp. 56–62, 2014.
- [16] J.-H. Ryu, D.-S. Kwon, and B. Hannaford, "Stable teleoperation with timedomain passivity control," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 365–373, 2004.
- [17] I. Nisky, F. A. Mussa-Ivaldi, and A. Karniel, "Analytical study of perceptual and motor transparency in bilateral teleoperation," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 6, pp. 570–582, 2013.
- [18] B. Siciliano and O. Khatib, Springer Handbook of Robotics. Springer, Jul. 2016, google-Books-ID: RTvADAAAQBAJ.
- [19] Y. Li and B. Hannaford, "Gaussian Process Regression for Sensorless Grip Force Estimation of Cable-Driven Elongated Surgical Instruments," *IEEE Robotics* and Automation Letters, vol. 2, no. 3, pp. 1312–1319, 2017.
- [20] M. M. Dalvand, B. Shirinzadeh, A. H. Shamdani, J. Smith, and Y. Zhong, "An actuated force feedbackenabled laparoscopic instrument for roboticassisted surgery," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 10, no. 1, pp. 11–21, 2014.
- [21] F. Anooshahpour, I. G. Polushin, and R. V. Patel, "Quasi-static modeling of the da Vinci instrument," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on.* IEEE, 2014, pp. 1308–1313.
- [22] A. I. A. Rivero, S. M. Alsaleh, J. K. Hahn, and A. Casals, "Towards Retrieving Force Feedback in Robotic-Assisted Surgery: A Supervised Neuro-Recurrent-Vision Approach," *IEEE transactions on haptics*, 2016.

- [23] A. M. Okamura and L. N. Verner, "Force & torque feedback vs force only feedback," in World Haptics Conference(WHC), 2009, pp. 406–410. [Online]. Available: doi.ieeecomputersociety.org/10.1109/WHC.2009.4810880
- [24] M. L. Yang, S. B. Schorr, I. Yan, and A. M. Okamura, "The effect of manipulator gripper stiffness on teleoperated task performance," in 2015 IEEE World Haptics Conference (WHC), Jun. 2015, pp. 494–499.
- [25] A. Talasaz, A. L. Trejos, and R. Patel, "The Role of Direct and Visual Force Feedback in Suturing using a 7-DOF Dual-Arm Teleoperated System," *IEEE transactions on haptics*, 2017.
- [26] T. L. Gibo, D. R. Deo, F. Q. Zhan, and A. M. Okamura, "Effect of load force feedback on grip force control during teleoperation: A preliminary study," 2014, pp. 379–383.
- [27] B. Weber and S. Schneider, "The Effects of Force Feedback on Surgical Task Performance: A Meta-Analytical Integration," in *Haptics: Neuroscience, Devices, Modeling, and Applications*, ser. Lecture Notes in Computer Science, M. Auvray and C. Duriez, Eds. Springer Berlin Heidelberg, 2014, pp. 150–157.
- B. Weber and C. Eichberger, "The Benefits of Haptic Feedback in Telesurgery and Other Teleoperation Systems: A Meta-Analysis," in Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being, M. Antona and C. Stephanidis, Eds. Springer International Publishing, 2015, pp. 394–405.
- [29] M. Mahvash, J. Gwilliam, R. Agarwal, B. Vagvolgyi, L.-M. Su, D. D. Yuh, and A. M. Okamura, "Force-feedback surgical teleoperator: Controller design and palpation experiments," in *Haptic interfaces for virtual environment and teleoperator systems, 2008. haptics 2008. symposium on.* IEEE, 2008, pp. 465– 471.

- [30] J. C. Gwilliam, M. Mahvash, B. Vagvolgyi, A. Vacharat, D. D. Yuh, and A. M. Okamura, "Effects of haptic and graphical force feedback on teleoperated palpation," in 2009 IEEE International Conference on Robotics and Automation, 2009, pp. 677–682.
- [31] L. Santos-Carreras, R. Beira, A. Sengl, R. Gassert, and H. Bleuler, "Influence of force and torque feedback on operator performance in a VR-based suturing task," *Applied Bionics and Biomechanics*, vol. 7, no. 3, pp. 217–230, Jan. 2010.
  [Online]. Available: https://content.iospress.com/articles/applied-bionics-and-biomechanics/abb503110
- [32] I. Nisky, A. M. Okamura, and M. H. Hsieh, "Effects of robotic manipulators on movements of novices and surgeons," *Surgical endoscopy*, vol. 28, no. 7, pp. 2145–2158, 2014. [Online]. Available: https://doi.org/10.1007/s00464-014-3446-5
- [33] I. Nisky, Y. Che, Z. F. Quek, M. Weber, M. H. Hsieh, and A. M. Okamura, "Teleoperated versus open needle driving: Kinematic analysis of experienced surgeons and novice users," in 2015 IEEE International Conference on Robotics and Automation (ICRA), May 2015, pp. 5371–5377.
- [34] K. Narazaki, D. Oleynikov, and N. Stergiou, "Robotic surgery training and performance," *Surgical Endoscopy And Other Interventional Techniques*, vol. 20, no. 1, pp. 96–103, 2006. [Online]. Available: https://doi.org/10.1007/s00464-005-3011-3
- [35] E. F. Hofstad, C. Vpenstad, M. K. Chmarra, T. Lang, E. Kuhry, and R. Mrvik, "A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance," *Surgical endoscopy*, vol. 27, no. 3, pp. 854–863, 2013. [Online]. Available: https://doi.org/10.1007/s00464-012-2524-9
- [36] T. N. Judkins, D. Oleynikov, and N. Stergiou, "Objective evaluation of expert and novice performance during robotic surgical training tasks,"

Surgical endoscopy, vol. 23, no. 3, p. 590, 2008. [Online]. Available: https://doi.org/10.1007/s00464-008-9933-9

- [37] C. D. Smith, T. M. Farrell, S. S. McNatt, and R. E. Metreveli, Assessing laparoscopic manipulative skills, 2001, vol. 181, no. 6. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0002961001006390
- [38] N. Nordin, S. Q. Xie, and B. Wnsche, "Assessment of movement quality in robot- assisted upper limb rehabilitation after stroke: a review," *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, p. 137, 2014. [Online]. Available: https://doi.org/10.1186/1743-0003-11-137
- [39] C. Richards, J. Rosen, B. Hannaford, C. Pellegrini, and M. Sinanan, "Skills evaluation in minimally invasive surgery using force/torque signatures," *Surgical endoscopy*, vol. 14, no. 9, pp. 791–798, 2000. [Online]. Available: https://doi.org/10.1007/s004640000230
- [40] Y. Sharon, T. S. Lendvay, and I. Nisky, "Instrument Orientation-Based Metrics for Surgical Skill Evaluation in Robot-Assisted and Open Needle Driving," arXiv preprint arXiv:1709.09452, 2017.
- [41] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of neural science*. McGraw-hill New York, 2000, vol. 4.
- [42] T. Flash, Y. Meirovitch, and A. Barliya, "Models of human movement: Trajectory planning and inverse kinematics studies," *Robotics and Autonomous Systems*, vol. 61, no. 4, pp. 330–339, 2013.
- [43] D. W. Franklin and D. M. Wolpert, Computational Mechanisms of Sensorimotor Control, 2011, vol. 72, no. 3. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0896627311008919
- [44] F. E. Pollick, U. Maoz, A. A. Handzel, P. J. Giblin, G. Sapiro, and T. Flash, "Three-dimensional arm movements at constant equi-

affine speed," Special Issue on "Cognitive Neuroscience of Drawing", vol. 45, no. 3, pp. 325–339, Mar. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S001094520800169X

- [45] Y. Sharon and I. Nisky, "Expertise, teleoperation, and task constraints affect the speed-curvature-torsion power law in RAMIS," *Journal of Medical Robotics Research*, 2018.
- [46] N. Hogan and D. Sternad, "Sensitivity of Smoothness Measures to Movement Duration, Amplitude, and Arrests," *Journal of motor behavior*, vol. 41, no. 6, pp. 529–534, 2009. [Online]. Available: https://doi.org/10.3200/35-09-004-RC
- [47] S. Balasubramanian, A. Melendez-Calderon, and E. Burdet, "A Robust and Sensitive Metric for Quantifying Movement Smoothness," *IEEE Transactions* on Biomedical Engineering, vol. 59, no. 8, pp. 2126–2136, 2012.
- [48] S. Pandey, M. D. Byrne, W. H. Jantscher, M. K. OMalley, and P. Agarwal, "Toward training surgeons with motion-based feedback: Initial validation of smoothness as a measure of motor learning," *Proceedings of the Human Factors* and Ergonomics Society Annual Meeting, vol. 61, no. 1, pp. 1531–1535, Sep. 2017. [Online]. Available: https://doi.org/10.1177/1541931213601747
- [49] S. Estrada, C. Duran, D. Schulz, J. Bismuth, M. D. Byrne, and M. K. OMalley, "Smoothness of surgical tool tip motion correlates to skill in endovascular tasks," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 5, pp. 647–659, Oct. 2016.
- [50] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da Vinci Surgical System," in 2014 IEEE International Conference on Robotics and Automation (ICRA), Jun. 2014, pp. 6434–6439.

- [51] Z. Chen, A. Deguet, R. H. Taylor, and P. Kazanzides, "Software Architecture of the Da Vinci Research Kit," in 2017 First IEEE International Conference on Robotic Computing (IRC), 2017, pp. 180–187.
- [52] G. Tholey, J. P. Desai, and A. E. Castellanos, "Force Feedback Plays a Significant Role in Minimally Invasive Surgery: Results and Analysis," *Annals of Surgery*, vol. 241, no. 1, pp. 102–109, 2005. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1356852/
- [53] J. Arata, H. Takahashi, S. Yasunaka, K. Onda, K. Tanaka, N. Sugita, K. Tanoue, K. Konishi, S. Ieiri, Y. Fujino, Y. Ueda, H. Fujimoto, M. Mitsuishi, and M. Hashizume, "Impact of network time-delay and force feedback on tele-surgery," *International Journal of Computer Assisted Radiology* and Surgery, vol. 3, no. 3, pp. 371–378, Sep. 2008. [Online]. Available: https://doi.org/10.1007/s11548-008-0228-3
- [54] M. E. Currie, A. Talasaz, R. Rayman, M. W. Chu, B. Kiaii, T. Peters, A. L. Trejos, and R. Patel, "The role of visual and direct force feedback in robotics-assisted mitral valve annuloplasty," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, no. 3, p. e1787, Sep. 2017. [Online]. Available: https://doi.org/10.1002/rcs.1787
- [55] L. N. Verner and A. M. Okamura, "Effects of Translational and Gripping Force Feedback are Decoupled in a 4-Degree-of-Freedom Telemanipulator," in Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07), Mar. 2007, pp. 286–291.
- [56] I. Nisky, M. H. Hsieh, and A. M. Okamura, "Uncontrolled Manifold Analysis of Arm Joint Angle Variability During Robotic Teleoperation and Freehand Movement of Surgeons and Novices," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 12, pp. 2869–2881, Dec. 2014.

- [57] J. Buzzi, E. D. Momi, and I. Nisky, "An Uncontrolled Manifold Analysis of Arm Joint Variability in Virtual Planar Position and Orientation Telemanipulation," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 2, pp. 391–402, Feb. 2019.
- [58] A. A. Faisal, L. P. J. Selen, and D. M. Wolpert, "Noise in the nervous system," *Nature Reviews Neuroscience*, vol. 9, no. 4, pp. 292–303, Apr. 2008. [Online]. Available: https://www.nature.com/articles/nrn2258
- [59] N. Stergiou and L. M. Decker, "Human movement variability, nonlinear dynamics, and pathology: Is there a connection?" *Human Movement Science*, vol. 30, no. 5, pp. 869–888, Oct. 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167945711000832
- [60] B. C. L. Touwen, "How normal is variable, or how variable is normal?" Early Human Development, vol. 34, no. 1, pp. 1–12, Sep. 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/037837829390035S
- [61] J. P. Kaipust, J. M. Huisinga, M. Filipi, and N. Stergiou, "Gait Variability Measures Reveal Differences Between Multiple Sclerosis Patients and Healthy Controls," *Motor Control*, vol. 16, no. 2, pp. 229–244, Apr. 2012. [Online]. Available: https://journals.humankinetics.com/doi/abs/10.1123/mcj.16.2.229
- [62] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature Neuroscience*, vol. 5, no. 11, pp. 1226–1235, Nov. 2002.
   [Online]. Available: https://www.nature.com/articles/nn963
- [63] H. Mller and D. Sternad, "Decomposition of Variability in the Execution of Goal-Oriented Tasks: Three Components of Skill Improvement," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 30, no. 1, pp. 212–233, Feb. 2004.

- [64] J. B. Dingwell, R. F. Smallwood, and J. P. Cusumano, "Trial-to-trial dynamics and learning in a generalized, redundant reaching task," *Journal of Neurophysiology*, vol. 109, no. 1, pp. 225–237, Oct. 2012. [Online]. Available: https://doi.org/10.1152/jn.00951.2011
- [65] C. G. Atkeson, C. H. An, and J. M. Hollerbach, "Estimation of Inertial Parameters of Manipulator Loads and Links," *The International Journal of Robotics Research*, vol. 5, no. 3, pp. 101–119, Sep. 1986. [Online]. Available: https://doi.org/10.1177/027836498600500306
- [66] Z. F. Quek, W. Provancher, and A. Okamura, "Evaluation of Skin Deformation Tactile Feedback for Teleoperated Surgical Tasks," *IEEE Transactions on Haptics*, pp. 1–1, 2018.
- [67] M. Mahvash and A. M. Okamura, "Enhancing Transparency of a Position-Exchange Teleoperator," in Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07), 2007, pp. 470–475.
- [68] J. K. Koehn and K. J. Kuchenbecker, "Surgeons and non-surgeons prefer haptic feedback of instrument vibrations during robotic surgery," *Surgical endoscopy*, vol. 29, no. 10, p. 2970, 2015.

## A Appendix: Full results

In this appendix we present the statistical results for the full task and for the extraction subtask. The statistical analysis is identical to the analysis that we present in the thesis.

In section A.1 we present the statistical analysis for the full task - including the insertion, repositioning and extraction. In section A.2 we present the statistical analysis for the extraction subtask. In each section, we present three figures corresponding to the learning contrast, aftereffect contrast, and final performance contrast. In addition, we present the results for the statistical tests.

### A.1 Results for the full task



Figure 12: Summary of the dVRK teleoperation *learning* during the full task. For each metric, median of difference between late teleoperation stage to early teleoperation stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.



Figure 13: Summary of the open needle driving *aftereffect* for the full task. For each metric, median of difference between early stage in the second open needle driving to late stage in the first open needle driving are presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.



Figure 14: Summary of the dVRK teleoperation *final performance* compared to baseline for the full task. For each metric, median of difference between late teleoperation stage to late first open stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.

Matria			Learning			Aftereffect		Fir	nal performa	nce
Metric name		NF	PE	DF	NF	PE	DF	NF	PE	DF
Fuit point oppor	р	0.375	0.0645	0.232	NaN	NaN	NaN	Final perfer           F         NF         PE           N         NaN         Nal           92         0.0195         0.10           601         -3.59         -3.0           N         NaN         Nal           92         0.00977         0.01           00         -1.94         -0.9'           92         0.00195         0.06           45         -1.42         -0.6'           32         0.275         0.37           44         4.83         18.4'           92         0.00195         0.001           0         -1610         -56           32         0.275         0.37           44         4.83         18.4'           92         0.00195         0.001           0         -1610         -56           22         0.131         0.01           5-05         0.0366         0.06           75         0.0137         0.69           5-06         3.74E-05         4.34E           22         0.00977         0.27           53         -3.31         -15           54         0.625	NaN	NaN
Exit point error	effect	0.504	0.953	1	NaN	NaN	NaN	NaN	NaN	NaN
Completion time	р	0.00195	0.00195	0.0195	0.232	0.0645	0.492	0.0195	Final performance       NF     PE       NaN     NaN       NaN     NaN       0.0195     0.105       -3.59     -3.06       NaN     NaN       0.00977     0.0137       -1.94     -0.979       0.00195     0.0645       -1.42     -0.663       0.275     0.375       4.83     18.6       0.00195     0.00195       -1610     -563       0.131     0.0137       0.0366     0.0641       0.0366     0.0641       0.0367     0.695       .74E-05     4.34E-07       0.00977     0.275       0.0363     -0.00137       0.625     0.00586       -3.31     -15       0.77     0.275       0.0385     -0.00438       0.0488     0.00391       0.168     0.166       0.322     1       0.0384     -0.0019       0.432     0.432       0.00195     -0.0019       0.432     0.432 <td>0.0371</td>	0.0371
Completion time	effect	6.15	10.3	10.9	-0.759	-1.37	-0.501	-3.59		-3.56
Total Eases nonnalized	р	0.00195	0.00391	0.00391	NaN	NaN	NaN	NaN	NaN	NaN
Total Force normalized	effect	0.825	0.975	1.43	NaN	NaN	NaN	Final perform           NF         PE           N         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           12         0.0195         0.105           01         -3.59         -3.06           N         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           0.00977         0.0137           06         -1.94         -0.979           02         0.00195         0.0645           45         -1.42         -0.663           32         0.275         0.375           4         4.83         18.6           92         0.00195         0.0019           0         -1610         -563           22         0.131         0.0137           6.05         0.0366         0.0642           75         0.0137         0.695           6.0625         0.00137           5.06         3.74E-05         1.39E-0           5.3         -3.31         -15           0.77         0.275           186 <td< td=""><td>NaN</td><td>NaN</td></td<>	NaN	NaN
Man fance 7 ania	р	1	0.846	1	0.375	0.131	DF         NF         PE           NaN         NaN         NaN         NaN           0.492         0.0195         0.105           -0.501         -3.59         -3.06           NaN         NaN         NaN           0.106         -1.94         -0.97           0.492         0.00195         0.0643           0         0.345         -1.42         -0.663           0.232         0.275         0.375           6.14         4.83         18.6           0.492         0.00195         0.00195           0.492         0.131         0.013           0.492         0.137         0.695           0.375         0.0137         0.695           0.375         0.0137         0.695           0.6         3.77E-06         3.74E-05         1.99E-           0.00586         0.625         0.0038           0.492         0.0488         0.0032	0.0137	0.432	
Max force - Z axis	effect	Learning         DF         NF           NF         PE         DF         NF           0.375         0.0645         0.232         NaN           0.504         0.953         1         NaN           0.00195         0.00195         0.0195         0.232           6.15         10.3         10.9         -0.759           0.00195         0.00391         0.00391         NaN           0.825         0.975         1.43         NaN           0.825         0.975         1.43         NaN           1         0.846         1         0.375           -0.118         0.0428         -0.0142         0.0392           0.922         0.492         0.322         0.131           -0.161         0.324         0.766         0.367           1         0.0645         0.131         0.432           -0.631         7.51         5.22         -0.981           0.084         0.00586         0.0195         0.375           745         1580         3180         -29.8           0.0371         0.0195         0.846         0.322           5.84E-06         -8.04E-06         8.24E-07	0.0392	-0.277	0.106	-1.94	-0.979	-0.264		
Man fance XV plane	р	0.922	0.492	0.322	0.131	1	0.492	0.00195	0.0645	0.492
Max force - A i plane	effect	-0.161	0.324	0.766	0.367	0.00279	0.345	-1.42	-0.663	-0.237
Man tanana 7 ania	р	1	0.0645	0.131	0.432	0.193	0.232	0.275	0.375	0.432
Max torque - Z axis	effect	-0.631	7.51	5.22	-0.981	17.7	6.14	4.83	18.6	9.85
DTW - force	р	0.084	0.00586	0.0195	0.375	0.193	0.492	0.00195	0.00195	0.00195
DIW-IOICE	effect	745	1580	3180	-29.8	-88.9	180	-1610	-563	-463
Path length	р	0.0371	0.0371	0.0195	0.846	0.846	0.922	0.131	0.0137	0.193
rath length	effect	0.0332	0.0522	0.0706	-0.0011	-0.00059	4.85E-05	F         NF         PE           N         NaN         NaN           N         NaN         NaN           92         0.0195         0.105           01         -3.59         -3.06           N         NaN         NaN           92         0.0195         0.105           01         -3.59         -3.06           N         NaN         NaN           N         NaN         NaN           N         NaN         NaN           0         0.00977         0.0137           06         -1.42         -0.663           32         0.275         0.375           14         4.83         18.6           92         0.00195         0.00195           0         -1610         -563           22         0.131         0.0137           50         -1610         -563           22         0.0366         0.0641           75         0.0137         0.695           2-06         3.74E-05         4.34E-07           22         0.00977         0.0137           E-06         2.11E-05         1.99E-03	0.0641	0.0146
Cincle domintion	р	0.232	0.322	0.846	0.322	0.432	0.375	NF         PE           NaN         NaN           0.00977         0.013           -1.42         -0.66           0.275         0.37           4.83         18.6           0.00195         0.001           -1610         -563           0.131         0.013           0.0366         0.064           0.0376         0.693           3.74E-05         4.34E           0.00977         0.013           3.74E-05         1.99E           3         0.625         0.004           0.0325         -0.004           0.0326         0.003           0.168         0.168           0.322         1           0.00896         -0.003           0.432         0.433           -0.00804         0	0.695	0.432
	effect	5.84E-06	-8.04E-06	8.24E-07	1.61E-05	-6.86E-06	3.77E-06	3.74E-05	4.34E-07	6.31E-06
Plana doriation	р	0.375	0.16	0.0273	0.557	0.77	0.922	0.00977	0.0137	0.0137
Fiane deviation	effect	4.75E-06	4.05E-06	9.53E-06	-2.19E-06	-2.73E-06	-1.18E-06	DF         NF         PE           NaN         NaN         NaN           NaN         NaN         NaN           A92         0.0195         0.100           0.501         -3.59         -3.00           NaN         NaN         NaN           1         0.00977         0.013           .106         -1.94         -0.97           .492         0.00195         0.064           .345         -1.42         -0.66           .232         0.275         0.37           .14         4.83         18.6           .492         0.00195         0.001           180         -1610         -563           .922         0.0137         0.69           .7E-06         3.74E-05         4.34E           .922         0.00977         0.013           .88E-06         2.11E-05         1.99E           .0553         -3.31         -15	1.99E-05	1.38E-05
Angular noth	р	1	0.084	0.432	0.492	0.557	0.00586	0.625	0.00586	1
Angular path	effect	0.72	8.17	-3.52	-6.19	4.27	5.53	-3.31	-15	0.44
DTW position	р	0.0195	0.00195	0.105	0.922	0.232	1	NaN         NaN           NaN         NaN           0.492         0.0195           -0.501         -3.59           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           1         0.00977           0.106         -1.94           0.492         0.00195           0.345         -1.42           0.232         0.275           6.14         4.83           0.492         0.00195           180         -1610           0.922         0.131           4.85E-05         0.0366           0.375         0.0137           3.74E-06         2.11E-05           0.922         0.00977           -1.18E-06         2.11E-05           0.00186         0.625           5.53         -3.31           1         0.77           0.00186         0.0835           0.492         0.0488           0.0302         0.168           1         0.322           0.00209         0.00896           0.846         0.432           0.00158 <td>0.275</td> <td>0.232</td>	0.275	0.232
DIW - position	effect	0.0741	0.0604	0.00135	0.0932	-0.0174	0.00186	0.0835	-0.00438	-0.0959
1/C nomen lans a	р	0.922	0.16	0.77	0.0645	0.557	0.492	0.0488	0.00391	0.00977
1/0 power law a	effect	-0.0256	-0.0423	-0.00181	-0.0812	-0.0394	0.0302	0.168	0.166	0.156
1/6 norman lans R	р	0.492	0.77	0.922	0.375	0.0488	1	0.322	1	0.625
1/0 power law $p$	effect	0.00519	0.00592	0.00303	-0.0102	-0.0142	0.00209	0.00896	0.375         0.432           18.6         9.85           0.00195         0.00195           -563         -463           0.0137         0.193           0.0641         0.0146           0.695         0.432           5         4.34E-07         6.31E-06           '         0.0137         0.0137           5         1.99E-05         1.38E-05           0.00586         1           -15         0.44           0.275         0.232           -0.00438         -0.0959           0.00391         0.00977           0.166         0.156           1         0.625           -0.0019         0.00681           0.432         0.922           -0.00446         -0.0004           0.00195         0.00195           1.94         1.67	
1/6 morrow love a	р	0.0488	0.084	0.695	0.375	0.275	0.846	0.432	0.432	0.922
1/0 power taw γ	effect	-0.0153	-0.0109	0.00698	-0.00496	-0.00993	1 $1$ $1$ $1$ NaN         NaN $1$ NaN         NaN $15$ 0.492         0.0195 $7$ -0.501         -3.59 $1$ NaN         NaN $1$ NaN         NaN $1$ 1         0.00977 $7$ 0.106         -1.94 $0.492$ 0.00195 $79$ 0.345         -1.42 $3$ 0.232         0.275 $7$ 6.14         4.83 $3$ 0.492         0.00195 $9$ 180         -1610 $6$ 0.922         0.131 $159$ 4.85E-05         0.0366 $2$ 0.375         0.0137 $-66$ 3.77E-06         3.74E-05 $7$ 0.922         0.00977 $-66$ 3.77E-06         3.74E-05 $7$ 0.0202         0.0325 $7$ 0.0586         0.625 $7$ 0.0492         0.0488 </td <td>0.00446</td> <td>-0.0004</td>	0.00446	-0.0004	
Iork	р	0.232	0.084	0.232	0.0488	0.432	0.492	0.00195	0.00195	0.00195
Jeik	effect	-0.158	-0.179	-0.154	0.186	0.0741	0.07	1.73	1.94	1.67

 ${\bf Table \ 3:}\ {\rm Statistical\ analysis}\ -\ {\rm Wilcoxon\ sign\ rank}\ -\ {\rm Full\ trial}$ 

Motnio nomo		Learning		Aftereffect			Final performance			
Metric name		NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF
	p model and $\chi^2$	p =0.	804, $\chi^2 =$	0.436	p =1	NaN, $\chi^2 =$	NaN	p =1	NaN, $\chi^2 =$	=NaN
Exit point error	р	0.842	0.829	1	NaN	NaN	NaN	NaN	NaN	NaN
	effect	-2.2	-2.3	-0.1	NaN	NaN	NaN	NaN	NaN	NaN
	p model and $\chi^2$	p =0	.581, $\chi^2 =$	1.09	p =0	.802, $\chi^2 =$	0.441	p =0	.917, $\chi^2 =$	=0.173
Completion time	р	0.842	0.551	0.88	0.815	0.997	0.855	0.913	0.991	0.958
	effect	-2.2	-4.1	-1.9	2.4	0.3	-2.1	-1.6	$\begin{tabular}{ c c c c c c } \hline $\mathbf{N}_{\mathrm{F}}, $\mathbf{N}_{\mathrm{T}}^{2} = $$ $$ $$ $$ $$ $$ $$ $$ $$ $$ $$ $$ $$	1.1
	p model and $\chi^2$	p =0	.112, $\chi^2 =$	4.38	p =1	NaN, $\chi^2 =$	NaN	p = l	NaN, $\chi^2 =$	=NaN
Total Force normalized	р	0.942	0.123	0.235	NaN	NaN	NaN	NaN	NaN	NaN
	effect	-1.3	-7.7	-6.4	NaN	NaN	NaN	NaN	al performand NF-DF 1 NaN, $\chi^2 = Na$ NaN $\lambda = 17, \chi^2 = 0.1$ 0.991 -0.5 NaN, $\chi^2 = 0.1$ 0.991 -0.5 NaN, $\chi^2 = 0.1$ 0.991 -0.5 NaN, $\chi^2 = 0.1$ $0.137, \chi^2 = 3.$ $0.208, \chi^2 = 3.$ $0.208, \chi^2 = 3.$ $0.208, \chi^2 = 3.$ $0.208, \chi^2 = 0.0$ 1 $0.137, \chi^2 = 3.$ $0.164, \chi^2 = 0.0$ 1 $0.137, \chi^2 = 3.$ 0.166 -7.2 $0.103, \chi^2 = 4.$ $0.038, \chi^2 = 4.$ $0.584, \chi^2 = 1.$ 0.583 3.1 $0.584, \chi^2 = 1.$ 0.987 -0.6 $0.882, \chi^2 = 0.3$ 0.933 1.4 $0.671, \chi^2 = 0.7$ 0.933 1.4 $0.671, \chi^2 = 0.7$ 0.987 0.66 $0.505, \chi^2 = 1.$ 0.829 -2.3 $= 0.45, \chi^2 = 1.$ 0.921 1.7	NaN
	p model and $\chi^2$	p =0.	907, $\chi^2 =$	0.196	p =0	$0.244, \chi^2 =$	2.82	p =0	$0.164, \chi^2 =$	=3.61
Max force - Z axis	р	0.999	0.913	0.933	0.224	0.842	0.519	0.268	0.195	0.983
	effect	0.2	1.6	1.4	6.5	2.2	-4.3	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	-0.7	
	p model and $\chi^2$	p =0.	625, $\chi^2 =$	0.939	p =0	.708, $\chi^2 =$	0.692	$\begin{tabular}{ c c c c } \hline Final performan \\ NF-PE NF-DF \\ \hline p = NaN, \chi^2 = NNaN NaNNaN NaNp = 0.917, \chi^2 = 0. \\ 0.913 0.991 \\ -1.6 -0.5 \\ \hline p = NaN, \chi^2 = N. \\ NaN NaN NaN \\ \hline p = 0.164, \chi^2 = 3 \\ 0.268 0.195 \\ -6.1 -6.8 \\ \hline p = 0.208, \chi^2 = 3 \\ 0.329 0.235 \\ -5.6 -6.4 \\ \hline p = 0.976, \chi^2 = 0. \\ 0.983 1 \\ -0.7 0.1 \\ \hline p = 0.137, \chi^2 = 3 \\ 0.246 0.16 \\ -6.3 -7.2 \\ \hline p = 0.103, \chi^2 = 4 \\ 0.384 0.711 \\ -5.2 3.1 \\ \hline p = 0.0475, \chi^2 = 0. \\ 0.0695 0.0989 \\ 8.7 8.1 \\ \hline p = 0.088, \chi^2 = 4 \\ 0.16 0.987 \\ 7.2 -0.6 \\ \hline p = 0.86, \chi^2 = 0.3 \\ 0.987 0.923 \\ -0.6 1.5 \\ \hline p = 0.882, \chi^2 = 0 \\ 0.987 0.923 \\ -0.6 1.5 \\ \hline p = 0.882, \chi^2 = 0 \\ 0.679 0.987 \\ 3.3 0.6 \\ \hline p = 0.55, \chi^2 = 1 \\ 0.46 -2.3 \\ \hline p = 0.45, \chi^2 = 1 \\ 0.695 0.902 \\ \hline \end{tabular}$	=3.14	
Max force - XY plane	р	0.913	0.599	0.842	0.695	0.965	0.842	0.329	0.235	0.977
	effect	-1.6	-3.8	-2.2	3.2	1	-2.2	-5.6	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	-0.8
	p model and $\chi^2$	p =0.	626, $\chi^2 =$	0.937	p =0	.865, $\chi^2 =$	0.289	Final performa           Final performa           2-DF         NF-PE         NF-DF           1 $p = NaN, \chi^2$ 1           ian         NaN         NaN           NaN         NaN         NaN           ian         NaN         NaN           ian         NaN         NaN           1 $p = 0.917, \chi^2 = 0.0200, \chi^2 = 0.0000$ 1           ian         NaN         NaN         NaN           ian         Science         0.164, $\chi^2 = 0.0000000000000000000000000000000000$	=0.049	
Max torque - Z axis	р	0.679	0.679	1	0.913	0.994	0.867	0.983	1	0.977
	effect	-3.3	-3.3	0	-1.6	0.4	2	-0.7	0.1	0.8
	p model and $\chi^2$	p =0	.217, $\chi^2 =$	3.05	p =0	0.195, $\chi^2 =$	3.27	$\begin{array}{c c} & p = 0.137, \ \chi^2 = 3 \\ \hline 0.246 & 0.16 \\ \hline -6.3 & -7.2 \\ \hline p = 0.103, \ \chi^2 = -6.103, \ \chi^2 = -6.103$	=3.97	
DTW - force	р	0.801	0.195	0.519	0.711	0.567	0.168	0.246	0.16	0.972
	effect	-2.5	-6.8	-4.3	3.1	-4	-7.1	-6.3	$\begin{array}{c} \operatorname{nal} \operatorname{performat}\\ \operatorname{NF-DF}\\ \hline \\ \operatorname{NaN}\\ \operatorname{NaN}\\ \hline \\ \operatorname{O.195}\\ -6.8\\ \hline \\ \operatorname{O.208}, \chi^2 = 1\\ \hline \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \operatorname{O.195}\\ -6.8\\ \hline \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \operatorname{O.235}\\ -6.4\\ \hline \\ \\ \end{array} \right)$	-0.9
	p model and $\chi^2$	p =0	.481, $\chi^2 =$	1.46	p =0	.951, $\chi^2 =$	0.101	p =0	$0.103, \chi^2$ :	=4.54
Path length	р	0.972	0.488	0.631	0.95	0.972	0.997	0.384	0.711	0.0882
	effect	-0.9	-4.5	-3.6	1.2	0.9	-0.3	-5.2	$\begin{array}{c} \text{inal performa} \\ \text{NF-DF} \\ = \text{NaN, } \chi^2 = 1\\ \hline \text{NaN} \\ \hline 0.917, \chi^2 = 0\\ \hline 0.991 \\ -0.5 \\ = \text{NaN, } \chi^2 = 1\\ \hline \text{NaN} \\ \hline 0.105 \\ -0.5 \\ \hline 0.105 \\ -0.68 \\ \hline 0.205 \\ -6.4 \\ \hline 0.205 \\ -7.2 \\ \hline 0.0987 \\ \hline 0.6 \\ \hline 0.993 \\ \hline 0.987 \\ -0.6 \\ \hline 0.993 \\ \hline 0.987 \\ -0.6 \\ \hline 0.993 \\ \hline 1.5 \\ \hline 0.882, \chi^2 = \\ \hline 0.983 \\ -1.5 \\ \hline 0.882, \chi^2 = \\ \hline 0.933 \\ -1.4 \\ \hline 0.6 \\ \hline 0.997 \\ 0.6 \\ \hline 0.829 \\ -2.3 \\ -2.3 \\ -2.3 \\ -2.3 \\ -2.3 \\ -2.3 \\ \hline 0.902 \\ 1.7 \\ \hline \end{array}$	8.3
	p model and $\chi^2$	p =0	.266, $\chi^2 =$	2.65	p =0	0.298, $\chi^2 =$	2.42	p =0	.0475, $\chi^2$	=6.09
Metric name         Exit point error         Completion time         Total Force normalized         Max force - Z axis         Max force - XY plane         Max torque - Z axis         DTW - force         Path length         Circle deviation         Plane deviation         Angular path         DTW - position         1/6 power law β         1/6 power law γ         Jerk	р	0.235	0.742	0.647	0.442	0.316	0.972	0.0695	0.0989	0.987
	effect	6.4	2.9	-3.5	4.8	5.7	0.9	8.7	al perform $NF-DF$ $NaN, \chi^2 = $	-0.6
	p model and $\chi^2$	p =0	.379, $\chi^2 =$	1.94	p =0.	973, $\chi^2 =$	0.0542	p =0	$0.584, \chi^2 =$	=1.08
Plane deviation	р	0.958	0.551	0.384	0.987	0.972	0.997	$\begin{array}{rrrr} \text{NF-PE} & \text{NF-DF} \\ \text{p} = \text{NaN}, \chi^2 \\ \text{NaN} & \text{NaN} \\ \text{NaN} & \text{NaN} \\ \text{p} = 0.917, \chi^2 \\ 0.913 & 0.991 \\ -1.6 & -0.5 \\ \text{p} = \text{NaN}, \chi^2 \\ 0.913 & 0.991 \\ -1.6 & -0.5 \\ \text{p} = \text{NaN}, \chi^2 \\ 0.268 & 0.195 \\ -6.1 & -6.8 \\ \text{p} = 0.208, \chi^2 \\ 0.268 & 0.195 \\ -6.1 & -6.8 \\ \text{p} = 0.208, \chi^2 \\ 0.268 & 0.195 \\ -6.1 & -6.8 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.329 & 0.235 \\ -5.6 & -6.4 \\ \text{p} = 0.976, \chi^2 \\ 0.983 & 1 \\ -0.7 & 0.1 \\ \text{p} = 0.103, \chi^2 \\ 0.984 & 0.72 \\ 0.695 & 0.9089 \\ \text{s}.7 & \text{s}.1 \\ \text{p} = 0.684, \chi^2 \\ 0.987 & 0.23 \\ 0.987 & 0.23 \\ 0.987 & 0.23 \\ 0.987 & 0.23 \\ 0.987 & 0.23 \\ 0.988 & 0.933 \\ 1.9 & 1.4 \\ \text{p} = 0.671, \chi^2 \\ 0.679 & 0.987 \\ 3.3 & 0.6 \\ \text{p} = 0.505, \chi^2 \\ 0.472 & 0.829 \\ -4.6 & -2.3 \\ \text{p} = 0.45, \chi^2 \\ 0.695 & 0.902 \\ -3.2 & 1.7 \end{array}$	0.583	0.726
	effect	1.1	-4.1	-5.2	-0.6	-0.9	-0.3	0.9	3.9	3
	p model and $\chi^2$	p =0	.167, $\chi^2 =$	3.58	p =	$0.301, \chi^2 =$	= 2.4	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	=4.86	
Angular path	р	0.342	0.913	0.168	0.695	0.268	0.742	0.16	0.987	0.117
	effect	-5.5	1.6	7.1	-3.2	-6.1	-2.9	7.2	-0.6	-7.8
	p model and $\chi^2$	p =0	.404, $\chi^2 =$	1.81	p =	0.73, $\chi^2 =$	0.63	p =0	0.86, $\chi^2 =$	0.302
DTW - position	р	0.902	0.384	0.647	0.786	0.999	0.757	0.987	0.923	0.855
	effect	-1.7	-5.2	-3.5	2.6	-0.2	-2.8	-0.6	$\begin{array}{c} \begin{array}{c} \begin{array}{c} 1.5 \\ p \\ \hline p \\ \hline p \\ \hline p \\ \hline n \\ N \\ NaN \\ \hline p \\ \hline 0.991 \\ 0.991$	2.1
	p model and $\chi^2$	p =0	$0.73, \chi^2 =$	0.63	p =0	$0.213, \chi^2 =$	3.09	p =0	$0.882, \chi^2 =$	=0.25
$1/6$ power law $\alpha$	р	0.786	0.999	0.757	0.726	0.186	0.583	0.88	0.933	0.991
	effect	2.6	-0.2	-2.8	-3	-6.9	-3.9	1.9	1.4	-0.5
	p model and $\chi^2$	p =0	.882, $\chi^2 =$	0.25	p =0	$0.334, \chi^2 =$	2.19	p =0	.671, $\chi^2$ =	=0.797
1/6 power law $\beta$	р	0.991	0.88	0.933	0.891	0.583	0.316	0.679	0.987	0.772
	effect	0.5	1.9	1.4	1.8	-3.9	-5.7	3.3	al performation of the second state of the se	-2.7
	p model and $\chi^2$	p =0	$.221, \chi^2 =$	3.02	p =0	.703, $\chi^2 =$	0.705	p =0	$0.505, \chi^2 =$	=1.37
Completion time Total Force normalized Max force - Z axis Max force - XY plane Max torque - Z axis DTW - force Path length Circle deviation Plane deviation Plane deviation 1/6 power law $\alpha$ 1/6 power law $\gamma$ Jerk	р	0.958	0.235	0.37	0.923	0.891	0.679	0.472	0.829	0.829
	effect	-1.1	-6.4	-5.3	1.5	-1.8	-3.3	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	-2.3	2.3
	p model and $\chi^2$	p =0.	951, $\chi^2 =$	0.101	p =0	.772, $\chi^2 =$	0.519	p =	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	=1.6
Jerk	р	0.997	0.972	0.95	0.772	0.855	0.987	0.695	0.902	0.427
	effect	0.3	-0.9	-1.2	2.7	2.1	-0.6	-3.2	1.7	4.9

### Table 4: Statistical analysis - Kruskal-Wallis test - Full trial



### A.2 Results for the extraction subtask

Figure 15: Summary of the dVRK teleoperation *learning* during the extraction subtask. For each metric, median of difference between late teleoperation stage to early teleoperation stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.



Figure 16: Summary of the open needle driving *aftereffect* for the extraction subtask. For each metric, median of difference between early stage in the second open needle driving to late stage in the first open needle driving are presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.



Figure 17: Summary of the dVRK teleoperation *final performance* compared to baseline for the extraction subtask. For each metric, median of difference between late teleoperation stage to late first open stage is presented. Each marker represent individual subject. Error bars shows median of each group difference (NF, PE, and DF) and 95% bootstrap confidence interval. Grey asterisks represent the result of a Wilcoxon sign rank test for difference from zero.

Matria nama			Learning			Aftereffect		Fii	nal performa	nce
Metric name		NF	$\mathbf{PE}$	DF	NF	PE	DF	NF	PE	DF
Fait a sint summ	р	0.375	0.0645	0.232	NaN	NaN	NaN	Final perference           NF         PE           N         NaN         NaN           N         NaN         NaN           5         0.492         0.84           44         -0.32         0.02           N         NaN         NaN           0.00391         0.001           3         -2.24         -1.2           2         0.0195         0.003           0.7         -0.828         -0.80           3         0.492         0.37           3         -0.422         1.43           0.00516         0.003           24         0.00516         0.002           5         0.557         0.02'           6         0.557         0.02'           6         0.557         0.002'           71         0.557         0.002'           71         0.557         0.002'           71         0.557	NaN	NaN
Exit point error	effect	0.504	0.953	1	NaN	NaN	NaN	NaN	Final performa           NF         PE           NaN         NaN           NaN         NaN           0.492         0.846           -0.32         0.027           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           0.00391         0.00195           -2.24         -1.21           0.0195         0.0391           -0.422         1.43           0.492         0.375           -0.422         1.43           0.00586         0.00391           -2.91         -123           0.375         0.0273           0.375         0.0273           0.0516         0.00195           8.96E-06         7.93E-06           0.557         0.105           0.557         0.105           0.557         0.105           0.0557         0.0371           0.557         0.0371           0.557         0.131           0.557         0.131           0.0223         0.131	NaN
Completion time	р	0.131	0.00195	0.00977	0.0645	0.00977	0.625	Final perform           NF         PE           NaN         NaN           NaN         NaN           0.492         0.846           -0.32         0.027           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           0.00391         0.00195           -2.24         -1.21           0.0195         0.00391           -0.828         -0.833           0.492         0.375           -0.422         1.43           0.00586         0.00391           -291         -123           0.00516         0.00391           -291         -123           0.00516         0.00195           6         8.96E-06           7.93E-00         -0.0195           6         -2.17E-06         2.95E-00           0         0.557         0.0273           5         -2.17E-06         2.95E-00           3         -0.0273         0.557           4         0.557         0.105           3         -0.0328         0.00201 <td>0.846</td> <td>0.77</td>	0.846	0.77
Completion time	effect	0.605	2.12	2.08	-0.494	-0.726	-0.144	-0.32	Final performar         PE         NaN         NaN         NaN         NaN         O.027         NaN         O.027         NaN         O.00195         4       -1.21         0.00391         8       -0.803         2       0.375         2       1.43         86       0.00391         16       0.00195         91       0.00195         92       0.0273         16       0.00916         91       0.00195         92       0.0273         16       0.00273         16       0.00195         93       0.557         3       2.85         7       0.105         28       0.00201         7       0.0371         92       0.131         82       0.0246         55       0.432         198       0.0104         95       0.0104	-0.115
Total Eason normalized	р	0.105	0.00977	0.00391	NaN	NaN	NaN	NaN	NaN	NaN
Total Force normalized	effect	0.113	0.224	0.307	NaN	NaN	t         Final perform           DF         NF         PE           NaN         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           0.625         0.492         0.846           -0.144         -0.32         0.027           NaN         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           0.0193         0.00193           0.113         -2.24         -1.21           0.232         0.0195         0.00391           -0.297         -0.828         -0.803           0.193         0.492         0.375           1.03         -0.422         1.43           1         0.00586         0.00391           0.275         0.375         0.0273           0.275         0.375         0.0273           0.0324         0.00516         0.00194           0.0375         0.00371         0.0273           1.40E-06         -2.17E-06         2.95E-0           0.846         0.557         0.105     <	NaN	NaN	
May force 7 avia	р	0.375	0.105	0.375	0.846	Altereffect           PE           N         NaN           N         NaN           445         0.00977         0           94         -0.726            N         NaN         2           N         NaN         2           N         NaN         2           N         NaN         2           M         NaN         2           Affectors         -0.338         0           22         0.492         0           35         -0.21         -1           20         0.557         0           31         0.557         0           33         0.492         0           34         0.77         0           35         0.557         0           36         0.00377         0           36         0.2077         1.4           7         0.322         0           60         -7.11E-07         1.4           7         0.322         0           82         -0.00173         -0           7         0.557         0           82         -0.003	1	0.00391	0.00195	0.105
Max Iorce - Z axis	effect	-0.513	Learning         Aftereffect           F         PE         DF         NF         PE           75         0.0645         0.232         NaN         NaN           04         0.953         1         NaN         NaN           31         0.00195         0.00977         0.0645         0.00977           05         2.12         2.08         -0.494         -0.726           05         0.00977         0.00391         NaN         NaN           13         0.224         0.307         NaN         NaN           75         0.105         0.375         0.846         0.084           13         -0.469         -0.267         -7.11E-05         -0.338           0.922         0.375         0.322         0.492           598         -0.103         0.16         -0.235         -0.21           32         0.625         1         0.922         0.557           23         -0.12         0.0869         0.158         0.277           .         0.0137         0.00391         0.131         0.557           .8         215         1510         -45.1         -10.9           05 <td< td=""><td>0.113</td><td>-2.24</td><td>-1.21</td><td>-0.889</td></td<>	0.113	-2.24	-1.21	-0.889			
May force XV plane	р	1	0.922	0.375	0.322	0.492	0.232	0.0195	0.00391	0.77
Max force - A r plane	effect	-0.0598	-0.103	0.16	-0.235	-0.21	-0.297	-0.828	-0.803	-0.0846
Max torque - Z axis	р	0.232	0.625	1	0.922	0.557	0.193	0.492	0.375	0.322
Max torque - Z axis	effect	-3.23	-0.12	0.0869	0.158	0.277	1.03	-0.422	1.43	2.87
DTW - force	р	1	0.0137	0.00391	0.131	0.557	1	0.00586	0.00391	0.131
DIW - Iorce	effect	94.8	215	1510	-45.1	-10.9	24	-291	-123	-61.3
Path length	р	0.105	0.846	0.0488	0.193	0.492	0.275	0.375	0.0273	0.0488
1 aun lengun	effect	-0.00535	0.00361	DF         NF           DF         NA           0.232         NaN           1         NaN           0.00977         0.0645         0           2.08         -0.494         0           0.00391         NaN         0           0.307         0.846         0           -0.267         -7.11E-05         0           0.375         0.322         0           0.0869         0.158         0           0.00391         0.131         1           1510         -45.1         0           0.0488         0.193         0           0.16         0.695         1           1.03E-06         4.35E-06         -2           0.557         0.77         1           1.09         <	0.00377	0.00324	0.00516	0.00916	0.0058	
Cincle deviation	р	0.432	0.375	0.16	0.695	0.557	0.375	0.00391	NF         PE           NaN         NaN           NaN         NaN           0.492         0.846           -0.32         0.027           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           NaN         NaN           0.00391         0.00195           -2.24         -1.21           0.0195         0.00391           -0.828         -0.803           0.492         0.375           -0.422         1.43           0.00586         0.00391           0.375         0.0273           0.375         0.0273           0.375         0.0273           0.375         0.0273           0.557         0.0273           0.557         0.105           0.0021         0.557           0.557         0.103           0.557         0.0371           0.0592         0.103           0.922         0.131           -0.00382         0.0246           0.625         0.432           0.00298	0.00391
	effect	5.71E-07	1.83E-07	1.03E-06	4.35E-06	-2.90E-06	-1.50E-06	8.96E-06		3.63E-06
Plana doriation	р	0.193	0.131	0.846	0.084	0.77	0.625	0.557	0.0273	0.922
I falle deviation	effect	-2.68E-06	7.23E-07	-8.05E-07	-3.33E-06	-7.11E-07	1.40E-06	F         NF         PE           NN         NaN         NaN           NN         NaN         NaN           NN         NaN         NaN           NN         NaN         NaN           NA         NaN         NaN           NA         NaN         NaN           225         0.492         0.846           144         -0.32         0.027           NN         NaN         NaN           NA         NaN         NaN           13         -2.24         -1.21           32         0.492         0.375           0.33         -0.422         1.43           1         -2.91         -123           275         0.375         0.027           324         0.00516         0.0091           5         15.3         2.	2.95E-06	1.18E-06
Angular path	р	0.131	0.00977	0.557	0.77	0.322	0.846	0.0273	0.557	0.232
Angular path	effect	9	10.9	1.39	-10.9	-4.54	1.5	15.3	2.85	18.9
DTW position	р	0.77	0.0273	0.00391	0.625	0.432	0.846	NF         PE           NaN         NaN           0.00391         0.00195           3         -2.24         -1.21           0         0.0195         0.00391           7         -0.828         -0.803           3         0.492         0.375           -0.422         1.43           0.00586         0.00391           -291         -123           5         0.375         0.0273           6         0.00516         0.00916           5         0.0557         0.0273           6         0.557         0.0273           6         0.557         0.105           6         0.557         0.103           8         -0.00428         0.00201           71         0.557         0.131	0.77	
D1W - position	effect	0.00143	0.0122	0.0155	-0.00159	-0.00173	Nan         Nan           N         NaN         Nan           977 $0.625$ $0.492$ '26 $-0.144$ $-0.32$ N         NaN         NaN           84         1 $0.00391$ :38 $0.113$ $-2.24$ 92 $0.232$ $0.0195$ 21 $-0.297$ $-0.828$ 57 $0.193$ $0.492$ 77 $1.03$ $-0.422$ 57         1 $0.00586$ $0.9$ $24$ $-291$ 92 $0.275$ $0.375$ $377$ $0.00324$ $0.00516$ 57 $0.375$ $0.00391$ E-06 $-1.50E-06$ $8.96E-06$ 77 $0.625$ $0.557$ E-07 $1.40E-06$ $-2.17E-06$ 22 $0.846$	0.00201	0.00152	
1/6 power law o	р	0.492	0.922	1	0.77	0.557	0.0371	0.557	0.0371	0.0273
	effect	-0.0683	-0.00676	-0.0126	0.0182	-0.0408	0.192	0.0592	0.103	0.212
1/6 power law B	р	0.322	0.432	0.16	0.557	0.625	0.0371	0.922	0.131	0.00977
	effect	0.0164	0.0151	0.0123	-0.00851	-0.00366	0.018	-0.00382	06         7.93E-06         3.63E-06         3.63E-06           7         0.0273         0.922           06         2.95E-06         1.18E-06           3         0.557         0.232           4         2.85         18.9           7         0.105         0.77           28         0.00201         0.00152           7         0.0371         0.0273           2         0.103         0.212           2         0.131         0.00977           82         0.0246         0.232           5         0.432         0.232	
1/6 power law $\gamma$	р	1	0.432	0.0488	0.625	0.625	0.0645	0.625	0.432	0.232
1/0 power raw /	effect	-0.00091	0.0134	0.0134	-0.00594	-0.00497	0.0081	NF         PE           NaN         NaN           NaN         NaN           NaN         NaN           0.492         0.846           -0.32         0.027           NaN         NaN           0.00391         0.0019           -2.24         -1.21           0.0195         0.0039           -0.422         1.43           0.00586         0.0039           -291         -123           0.375         0.0273           0.00516         0.0091           0.00516         0.0091           6         8.96E-06         7.93E-1           0.557         0.0273         0.557           0.557         0.0273         0.557           15.3         2.85         0.557           0.557         0.0273         0.557           0.557         0.037         0.0557           0.557         0.037         0.557           0.0557<	0.0104	0.013
Iork	р	0.0488	0.131	0.084	0.275	0.432	1	0.00195	0.00195	0.00391
JCIK	effect	-0.382	-0.24	-0.249	0.169	0.119	-0.032	Nin         Per           NAN         NaN           NaN         NaN           0.492         0.846           -0.32         0.027           NaN         NaN           0.00391         0.00195           -2.24         -1.21           0.0195         0.0391           -0.422         1.43           0.492         0.375           0.492         0.375           0.492         1.43           0.0556         0.00391           0.375         0.0273           0.375         0.0273           0.375         0.0273           0.557         0.0273           0.557         0.1051           0.557         0.1051           0.557         0.0371           0.557         0.0371           0.557         0.1031           0.557         0.131           0.0592         0.131           0.0592         0.131<	2	1.6

 ${\bf Table \ 5:} \ {\rm Statistical \ analysis \ - \ Wilcoxon \ sign \ rank \ - \ Extraction \ subtask}$ 

Motnio nomo			Learning			Aftereffect	;	Fin	al perform	nance
Metric name		NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF	NF-PE	NF-DF	PE-DF
	p model and $\chi^2$	p =0.	.804, $\chi^2 =$	0.436	p =1	NaN, $\chi^2 =$	NaN	p =1	NaN, $\chi^2$ =	=NaN
Exit point error	р	0.842	0.829	1	NaN	NaN	NaN	NaN	NaN	NaN
	effect	-2.2	-2.3	-0.1	NaN	NaN	NaN	NaN	NaN	NaN
	p model and $\chi^2$	p =0	$1.119, \chi^2 =$	4.25	p =0	$0.581, \chi^2 =$	: 1.09	p =0	.798, $\chi^2$ =	=0.452
Completion time	р	0.488	0.0989	0.631	0.842	0.88	0.551	0.867	0.991	0.801
	effect	-4.5	-8.1	-3.6	2.2	-1.9	-4.1	-2	0.5	2.5
	p model and $\chi^2$	p =0	$104, \chi^2 =$	4.53	p =1	NaN, $\chi^2 =$	NaN	p =1	NaN, $\chi^2$ =	=NaN
Total Force normalized	р	0.786	0.0934	0.329	NaN	NaN	NaN	NaN	NaN	NaN
	effect	-2.6	-8.2	-5.6	NaN	NaN	NaN	NaN	$\begin{array}{c} \hline 1 \text{ performan} \\ \hline NF-DF \\ \hline NN, \chi^2 = N \\ \hline NaN \\ \hline 798, \chi^2 = 0 \\ \hline 0.991 \\ \hline 0.5 \\ \hline 0.991 \\ \hline 0.105 \\ \hline 0.985 \\ \hline 0.105 \\ \hline 0.987 \\ \hline 0.105 \\ \hline 0.987 \\ \hline $	NaN
	p model and $\chi^2$	p =0.	.924, $\chi^2 =$	0.157	p =0	0.167, $\chi^2 =$	3.58	p =0	.0794, $\chi^2$	=5.07
Max force - Z axis	р	1	0.942	0.933	0.224	1	0.235	0.152	0.105	0.983
	effect	0.1	-1.3	-1.4	6.5	0.1	-6.4	-7.3	$\begin{array}{r c c c c c c c c c c c c c c c c c c c$	-0.7
	p model and $\chi^2$	p =0.	.703, $\chi^2 =$	0.705	p =0.	991, $\chi^2 =$	0.0181	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	=4.48	
Max force - XY plane	р	0.891	0.679	0.923	0.991	0.994	1	0.987	0.137	0.186
	effect	-1.8	-3.3	-1.5	0.5	0.4	-0.1	-0.6	al performan           NF-DF           NaN, $\chi^2 = N$ NaN           0.798, $\chi^2 = 0$ 0.991           0.5           NaN           NaN           NaN           0.0794, $\chi^2 = N$ NaN           0.0794, $\chi^2 = N$ 0.105	-6.9
	p model and $\chi^2$	p =0	$0.557, \chi^2 =$	1.17	p =0	$0.389, \chi^2 =$	1.89	p =0	.985, $\chi^2$ =	=0.031
Max torque - Z axis	р	0.535	0.772	0.923	0.815	0.726	0.356	0.987	0.987	1
	effect	-4.2	-2.7	1.5	2.4	-3	-5.4	-0.6	-0.6	0
	p model and $\chi^2$	p =0.	.0372, $\chi^2$ =	= 6.58	p =0	$0.374, \chi^2 =$	1.97	p =0	).486, $\chi^2$	=1.44
DTW - force	р	0.384	0.0278	0.427	0.695	0.342	0.829	0.88	0.457	0.757
	effect	-5.2	-10.1	-4.9	-3.2	-5.5	-2.3	-1.9	al performal NF-DF NaN, $\chi^2 = N$ NaN NaN 0.798, $\chi^2 = 0$ 0.991 0.5 NaN, $\chi^2 = N$ NaN NaN 0.0794, $\chi^2 = 0$ 0.105 	-2.8
	p model and $\chi^2$	p =0	$0.0247, \chi^2$	= 7.4	p =0	.935, $\chi^2 =$	0.134	p =0	.781, $\chi^2$ =	=0.495
Path length	р	0.427	0.018	0.304	0.994	0.965	0.933	0.815	1	0.815
	effect	-4.9	-10.7	-5.8	-0.4	1	1.4	-2.4	al performa NF-DF NaN, $\chi^2 = 1$ NaN 0.798, $\chi^2 = 0$ 0.991 0.5 NaN NaN 0.0794, $\chi^2 = 0$ 0.105 	2.4
	p model and $\chi^2$	p =0.	.798, $\chi^2 =$	0.452	p =0	$0.379, \chi^2 =$	: 1.94	p =0	).196, $\chi^2$	=3.25
Circle deviation	р	0.991	0.867	0.801	0.384	0.551	PE-DF         NF-PE           NaN $p = Na$ NaN $p = Na$ NaN         NaN           1.09 $p = 0.7$ $0.551$ $0.867$ -4.1         -2           NaN $p = 0.7$ $0.551$ $0.867$ -4.1         -2           NaN $p = 0.7$ $0.358$ $p = 0.0^\circ$ $0.235$ $0.152$ -6.4         -7.3 $0.0181$ $p = 0.1$ $0.987$ -0.6 $1.89$ $p = 0.9$ $0.356$ $0.987$ -5.4         -0.6 $1.97$ $p = 0.4$ $0.829$ $0.88$ -2.3         -1.9 $0.134$ $p = 0.7$ $0.933$ $0.815$ $1.4$ -2.4 $1.94$ $p = 0.1$ $0.958$ 1 $-1.1$ -0.1 $0.942$ $0.13$ $0.766$ $p = 0.3$ $0.75$	0.268	0.257	
	effect	0.5	-2	-2.5	5.2	4.1	-1.1	-0.1	6.1	6.2
	p model and $\chi^2$	p =0	$.298, \chi^2 =$	2.42	p =	$0.42, \chi^2 =$	1.74	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	=4.35	
Plane deviation	р	0.268	0.786	0.647	0.615	0.412	0.942	0.13	0.958	0.224
	effect	-6.1	-2.6	3.5	-3.7	-5	-1.3	-7.6	-1.1	6.5
	p model and $\chi^2$	p =0.	.0979, $\chi^2$ =	= 4.65	p =0	.703, $\chi^2 =$	0.705	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	=1.88	
Angular path	р	1	0.145	0.152	0.923	0.891	0.679	0.427	0.991	0.503
	effect	0.1	7.4	7.3	1.5	-1.8	-3.3	4.9	0.5	-4.4
	p model and $\chi^2$	p =0	$1.114, \chi^2 =$	4.35	p =0	$0.756, \chi^2 =$	0.56	p =0	).324, $\chi^2$	=2.25
DTW - position	p	0.503	0.0934	0.599	1	0.801	0.786	0.316	0.535	0.923
	effect	-4.4	-8.2	-3.8	0.1	-2.5	-2.6	-5.7	$\begin{array}{cccccccccccccccccccccccccccccccccccc$	1.5
	p model and $\chi^2$	p =0.	$.917, \chi^2 =$	0.173	p =0	$0.099, \chi^2 =$	4.62	p =0	$.745, \chi^2 =$	=0.588
$1/6$ power law $\alpha$	p	0.913	0.991	0.958	0.913	0.235	0.105	0.95	0.891	0.726
	effect	-1.6	-0.5	1.1	1.6	-6.4	-8	1.2	-1.8	-3
	p model and $\chi^2$	p =0.	.921, $\chi^2 =$	0.165	p =	$0.18, \chi^2 =$	3.42	p =0	.716, $\chi^2$ =	=0.668
$1/6$ power law $\beta$	p	0.977	0.913	0.977	0.994	0.224	0.268	0.829	0.711	0.977
	effect 2	0.8	1.6	0.8	-0.4	-6.5	-6.1	-2.3	-3.1	-0.8
1/6	p model and $\chi^2$	p =0	$J.44, \chi^- =$	1.64	p =0	$0.384, \chi^2 =$	0.525	p =0	J.823, χ <sup>2</sup>	=0.39
$1/\sigma$ power law $\gamma$	p	0.647	0.427	0.933	0.972	0.398	0.535	Final p           2-DF         NF-PE         N           NaN         NaN         NaN           NaN         NaN         NaN           NaN         NaN         NaN           S551         0.867         0 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ -2         -2 $(4.1)$ $(-2, -7, -3)$ -2 $(3.3)$ 0.987         0 $(0.1)$ -0.6         -0 $(0.1)$ -0.6         -0 $(2.3)$ -1.9         -1.4 $(2.3)$ -1.9         -1.4 $(2.3)$ -1.9         -1.4 $(2.3)$ -1.9         -1.1 $(2.6)$ -5.7         -2 $(2.6)$ -5.7         -2 $(2.6)$ -5.7         -2 $(2.6)$	0.88	0.994
	n model and - 2	-3.0	-4.9 650 x <sup>2</sup>	-1.4	-0.9	-0.1	-4.4	-2.3	-1.9	-1.62
Iork	p model and $\chi^-$	p =0.	$\frac{1002}{0.711}$	0.804	p =0	0.842	0.027	p =0	Pie NF-DF P = NA, $\chi^2$ = N NaN p =0.798, $\chi^2$ = 67 0.991 2 0.5 p =NA, $\chi^2$ = 67 0.994, $\chi^2$ = 67 0.994, $\chi^2$ = 10 NaN p =0.0794, $\chi^2$ = 10 NaN p =0.085, $\chi^2$ = 10 NaN p =0.196, $\chi^2$ = 10 NaN p =0.24, $\chi^2$ = 10 NaN p =0.711 3 -3.1 p =0.823, $\chi^2$ 29 0.88 3 -1.9 p =0.443, $\chi^2$ 03 1 4 -0.1	0 510
JUIK	P	0.095	2.1	0.1	1.933	0.042	0.911	0.503	0.1	4.9
	enect	-3.4	-0.1	0.1	1.4	4.4	0.0	-4.4	-0.1	4.0

### Table 6: Statistical analysis - Kruskal-Wallis test - Extraction subtask

## **B** Appendix: Jerk metrics analysis

Jerk is the 3rd derivative of the movement, and is used as a measure for movement smoothness. The lower the jerk, the smoother the movement. There are several suggested methods to calculate the jerk of a movement [46]. In this appendix, we present an analysis on the correlation between the different methods to calculate the jerk to completion time and path length. We analyze the methods presented in [46].

We use the data of the insertion subtask recorded in the teleoperation setup of the experiment. Each data point includes the calculated jerk, path length, and completion time of each insertion subtask movement, total of 1800 data points (60 trials for each one of the 30 participants). We calculate the jerk with 6 different methods: (1) root mean squared jerk, (2) dimensionless jerk, (3) integrated squared jerk, (4) mean squared jerk normalized by peak speed, (5) integrated absolute jerk, and (6) mean squared absolute jerk normalized by peak speed.

For each method to calculate the jerk, we present a figure that includes two graphs - the jerk of the movement vs completion time and the jerk of the movement vs path length. We present the correlation between the jerk and time and between the jerk and path length for each feedback condition (no feedback - NF, position exchange -PE, and direct feedback - DF).



Figure 18: Results for root mean squared jerk



Figure 19: Results for dimensionless jerk



Figure 20: Results for integrated squared jerk



Figure 21: Results for mean squared jerk normalized by peak speed



Figure 22: Results for integrated absolute jerk



Figure 23: Results for mean absolute jerk normalized by peak speed

In this study, we use the root mean squared jerk method. This method normalizes the jerk by movement duration, and has units of jerk. As a result, this jerk metric
is less correlated with completion time. Because the jerk metric spans over several orders of magnitudes, we present and analyze the logarithm of the metric.

## תקציר

מערכות ניתוח רובוטי זעיר פולשני מספקות יתרונות רבים למנתח ולמנותח בהשוואה לניתוח פתוח וניתוח לפרוסקופי. יחד עם זאת, מתן משוב כוח, אשר מהווה חלק חיוני בהצלחת פרוצדורות ניתוחיות רבות, מהווה אתגר לא פתור ומרבית המערכות הקליניות לניתוח רובוטי זעיר פולשני אינן מספקות משוב כוח. הבנת ההשפעה של משוב כוח על הביצועים והלמידה של המנתח יכולה להיות שימושית בפיתוח של אלגוריתמים למתן משוב כוח ולבקרה של מערכות הפעלה מרחוק. במחקר זה בחנו את הביצועים של נבדקים לא מנוסים תחת מצבים שונים של משוב כוח במטלה של החדרת מחט ניתוחית לאובייקט הומוגני רך – רקמה מלאכותית. בנינו מערכת ניסוי לאפיון מסלולי התנועה של הנבדקים והכוחות שהנבדקים הפעילו על הרקמה המלאכותית. הנבדקים ביצעו תחילה את המטלה במצב פתוח, עם מחזיק מחט ניתוחי רגיל, לאחר מכן ביצעו את המטלה בהפעלה מרחוק באחד משלושה תנאים של משוב כוח: (1) ללא משוב כוח, (2) משוב כוח מבוסס על שגיאת מיקום הרובוט (Position exchange), (3) ומשוב כוח שהוקלט באופן ישיר מחיישן כוח, לבסוף חזרו לבצע את המטלה עם מחזיק המחט במצב פתוח. על מנת לכמת את ההשפעה של המצבים השונים של משוב הכוח על איכות ביצוע מטלת החדרת המחט, פיתחנו מטריקות חדשות על מנת לאמוד את הקינמטיקה של החדרת המחט ואת כוחות האינטראקציה עם הרקמה. בנוסף למטריקות שפיתחנו, השתמשנו גם במטריקות קלאסיות כגון: הצלחה במשימה וזמן סיום המטלה. עבור כל מצב של משוב כוח, ניתחנו את הביצועים הסופיים במערכת ההפעלה מרחוק, את השיפור במהלך ההפעלה מרחוק, ואת ההשפעה של השימוש במערכת הפעלה מרחוק על ביצוע המטלה במצב פתוח באמצעות מחזיק מחט. מצאנו כי אין הבדל מובהק בין 3 הקבוצות בביצועים הסופיים ובהשפעה של הפעלה מרחוק על המצב הפתוח. בחלק מהמטריקות שבחנו, רק שני המצבים בהם הנבדקים קיבלו משוב כוח הציגו שיפור מובהק סטטיסטית במהלך ההפעלה מרחוק, אך כאשר ביצענו השוואה בין השיפור של 3 הקבוצות של מתן משוב הכוח, אף אחד מהאפקטים לא הגיע למובהקות סטטיסטית. אנחנו דנים בסיבות אפשריות לביצועים הדומים יחסית בין מצבי משוב הכוח. לסיכום, פיתחנו מספר מטריקות חדשות לבחינת איכות ביצוע החדרת מחט ניתוחית, אך גם עם המטריקות החדשות, היתרון של שיטות עדכניות למתן משוב כוח למטלה שדורשת אינטראקציה עם רקמה הומוגנית רכה מוטל בספק.

אוניברסיטת בן גוריון בנגב הפקולטה למדעי ההנדסה המחלקה להנדסה ביורפואית

## ניתוח מכוון מנתח של החדרת מחט בניתוח רובוטי זעיר פולשני תחת מימושים שונים של משוב כוח

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת: לידור בכר מנחה: ד"ר אילנה ניסקי

תאריך: 17.05.19

26/06/19\_\_\_\_\_

תאריך: 19/7/2019

אוניברסיטת בן גוריון בנגב הפקולטה למדעי ההנדסה המחלקה להנדסה ביורפואית

## ניתוח מכוון מנתח של החדרת מחט בניתוח רובוטי זעיר פולשני תחת מימושים שונים של משוב כוח

חיבור זה מהווה חלק מהדרישות לקבלת תואר מגיסטר בהנדסה

מאת: לידור בכר

מנחה: ד"ר אילנה ניסקי

אדר א' תשע"ט

פברואר 2019