Multi-sensory Guidance and Feedback for Simulation-based Training in Robot Assisted Surgery: a Preliminary Comparison of Visual, Haptic, and Visuo-Haptic

Guido Caccianiga*[1], Andrea Mariani*[2], *Student Member, IEEE*, Chiara Galli de Paratesi*[3], Arianna Menciassi[2], *Senior Member, IEEE*, and Elena De Momi[3], *Senior Member, IEEE*

Abstract— **Nowadays, robot assisted surgery training relies more and more on computer-based simulation. However, the application of such training technologies is still limited to the early stages of practical training. To broaden the usefulness of simulators, multi-sensory feedback augmentation has been recently investigated. This study aims at combining initial predictive (guidance) and subsequent error-based (feedback) training augmentation in the visual and haptic domain. 32 participants performed 30 repetitions of a virtual reality task resembling needle-driving by using the surgeon console of the da Vinci Research Kit. These trainees were randomly and equally divided into four groups: one group had no training augmentation, while the other groups underwent visual, haptic and visuo-haptic augmentation, respectively. Results showed a significant improvement, initially introduced by guidance, in the task completion capabilities of all the experimental groups against control. In terms of accuracy, the experimental groups outperformed the control group at the end of training. Specifically, visual guidance and haptic feedback played a significant role in error reduction. Further investigations on long term learning could better delineate the optimal combination of guidance and feedback in these sensory domains.**

I. Introduction

A. Robot Assisted Surgery

Since their early introduction in medicine, robots have gradually started to become ubiquitous in rehabilitation and several surgical fields [1].

Robot-Assisted Minimally Invasive Surgery (RAMIS) is an example of how the clinical advantages of laparoscopy have been integrated with the technical improvements introduced by robotics. The da Vinci Surgical System (by Intuitive Surgical Inc., Sunnyvale, CA) is currently the most popular surgical robot in RAMIS. It is a teleoperated system where the surgeon manipulates a couple of robotic masters while seating at a console located away from the patient.

Despite the upgrades introduced by such a robotic architecture, the complexity of surgery as a sensorimotor task, together with the need for learning how to deal with these new control modalities [2], imply the distinct need for efficient training in robotic surgery [3].

*G.C., A.M., and C.G.d.P. equally contributed to this work.

[1]Guido Caccianiga was with the Department of Electronics, Information, and Bioengineering, Politecnico di Milano, IT, at the time of the study. Now he is with the Haptics Intelligence Department, Max-Planck-Institute for Intelligent Systems, Stuttgart, DE. caccianiga@is.mpg.de

[2]Andrea Mariani and Arianna Menciassi are with The BioRobotics Institute and the Department of Excellence in Robotics & AI, Scuola Superiore Sant'Anna, Pisa, IT.

[3]Chiara Galli de Paratesi and Elena De Momi are with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, IT.

B. Surgical Simulation and Training

In the last decades, simulation-based education has come to the forefront for the development of the practical surgical skills [4]. As of today, the vast majority of departments of surgery delivers simulation-based training to their trainees on daily basis [5]. Simulators for RAMIS training range from inanimate box trainers and manikins, to virtual reality (VR) platforms. These latter have been proven to enhance skill learning [6]. However, the use of such VR training seems to be limited to surgical residents and specific to the early stages of the training curricula [7].

To push forward the adoption of training simulators across different surgical fields and proficiency levels, various types of feedback augmentation have been attempted and tested [8].

Moreover, virtual environments open the appealing opportunity to recreate a variety of different settings which can be easily modified in real-time, as well as to extract an ample amount of simulation parameters and metrics. Together with the exploitation of these advantages to perform quantitative performance assessment [9], augmented feedback represents a further application of these available data to enhance motor learning.

C. Augmented Feedback

Feedback has been long recognized as a central component for the education, motivation and support of the trainee towards proficiency [10]. It is in fact thought to have a guiding role when used by the learner to produce more accurate performance on subsequent trials [11].

Augmented feedback, or extrinsic feedback, specifically refers to a piece of information that cannot be elaborated without an external source, and which is provided to the trainee by a trainer or a display [12]. The word display does not refer only to visual devices but in general to any feedback interface in multiple sensory domains. Computer-based VR platforms allow to perform this elaboration by taking as an input the simulation parameters and metrics, as well as they offer the possibility to visually display the resulting augmented feedback in the simulation environment.

Moreover, the combination of robotics with simulation as in robotic surgery simulation-based training further increases the amount of available data (e.g. robot kinematics and dynamics) and the way to display feedback. In fact, robots can be regarded as haptic displays.

This study fits into the context of simulation-based training for robotic surgery with the aim of optimizing augmentation by exploiting intrinsic properties of simulation and robotics to address multiple sensory domains.

II. Related Work

A. Augmentations in Motor Learning

Before addressing the literature related to augmentation in multiple sensory domains, it is useful to define also the concept of guidance which is strictly related to motor learning, training, and augmentation. According to [13], guidance can be described as any kind of physical or verbal assistance to prevent the user from committing methodological and/or large amplitude errors. In the following sections, we will refer to guidance or assistance with the specific meaning of active (or predictive) methods for sensory augmentation. On the other side, feedback will be regarded as the sensory augmentation which is exclusive consequence of the user's action (e.g. their error).

According to Zhu et al., the choice of the type of augmentation should be related to the task [14]. Visual augmentation is mainly connected to hand-eye coordination, context awareness, depth perception, or methodology; haptic augmentation is mainly involved in force learning, manipulation, and tactile discrimination; the auditory augmentation is mostly linked to strongly time-related and repetitive movements.

B. Visual Augmentation

Visual feedback was initially investigated in rehabilitation. It was demonstrated that visual cues might help in maintaining a predetermined trajectory by making corrective adjustments whenever the feedback indicates a deviation from the desired path. This can help, for example, during the traditional arm reaching movements [15].

Several studies also applied visual feedback in surgical training to investigate its impact over performance. Reiley et al. displayed haptic interaction forces as graphical overlays on the master console of a da Vinci surgical system [16]; similarly, Judkins et al. implemented an Augmented Reality environment by introducing visual cues in laparoscopic images to inform the subject in real-time about the magnitude of the exerted gripping forces [17]. In both cases, the developed visual force feedback improved the performance on trainees. Simulators such as the da Vinci Skill Simulator (dVSS) mostly integrate basic forms of visual guidance to teach how to control the robotic platform or to perform simple object manipulation [18].

Malpani et al. developed an advanced virtual coaching framework able to deliver real-time visual augmentation during suture on the dVSS; the visual overlays aimed at providing both guidance and feedback throughout the most challenging phases of the suturing task. Even though the visual cues were mostly found as intuitive, the overall performance did not improve significantly possibly due to inefficacy of the user study design [19].

C. Haptic Augmentation

Haptic augmentation has proven to be effective in many surgery-related training experiments; it is thought to improve fidelity, realism, and training efficacy [20] especially for novices or resident surgeons [21].

Shahbazi et al. proposed an Expert In the Loop (EIL) system to incorporate direct haptic guidance (also referred to as *virtual fixture* or *active constraint*) on da Vinci training. Their argument was that, in teleoperated systems, visual guidance might not be as effective as haptic guidance since it only shows where or how to move the robotic tool tip, but not how to generate such a movement through the master manipulator [22].

Furthermore, Enayati et al. modulated haptic assistance on a VR training task using the da Vinci Research Kit (dVRK [23]) according to the trainee's performance. Results showed promising outcomes in terms of faster task completion time [24]. This is a well known concept in rehabilitation (usually referred to as *guidance hypothesis*), where the amount of assistance is consequently mostly delivered "as needed". Nevertheless, the benefits of haptic guidance for skill learning and retention is still debated and different studies agree that this modality can considerably alter the dynamic of the task and even impede motor learning [25] [26].

Scheeler et al. found that haptic guidance is more suited for improving movement efficiency, while haptic feedback for improving its effectiveness; the authors suggest that the two modalities can be additionally deployed on different degrees of freedom [27].

Most of the RAMIS training experiments have been carried out deploying alternative haptic feedback displays such as wrist squeezing devices [28], voice coils [29], tactile actuators [30], or mixed approaches [31] to render the measured or simulated haptic clues. The dVRK is a well-known research platform in robotic surgery that allows to apply 6 DoF force-torque at its master manipulators [23]; this study fits in this context together with other groups that have been exploring the dVRK stability boundaries and capability as haptic display for surgical simulation [22], [32], [33].

D. Auditory Augmentation

Auditory augmentation has proven to be particularly effective in optimizing cyclic movements since the human auditory system is very sensitive in perceiving rhythms and time-dependent variations; this could be the main reason why it is more suited for sport training and rehabilitation than surgery [34]. Surgeons and novices found, in fact, vibrotactile feedback more useful than auditory during da Vinci dry lab training. Most of them also approved the combination of the two modalities [35]. Caccianiga et al. jointly deployed auditory and visual feedback in a RAMIS training scenario; visual cues were used to boost accuracy while sound was used to impose specific temporal constraints to task execution without overloading the visual domain [36].

E. Mixed Sensory Augmentation

To the best of our knowledge, a direct comparison or optimized fusion of the aforementioned augmentation modalities has not been presented yet in the context of robot-assisted surgical training; anyway, useful results can be extracted from other research contexts.

For example, Yoon et al. found that haptic assistance reduced task completion time and effort with respect to visual assistance on a mobile robot driving task [37]. Marchal-Crespo et al. used instead a tennis simulator to compare visual and haptic domain; haptic guidance was found especially suitable for less-skilled subjects, while visual feedback gave higher benefit to highly-skilled participants [38].

Sigrist et al. extensively tested a VR rowing simulator over different feedback domains and the study revealed that a well-designed concurrent (i.e., during the task) feedback can lead to complex task learning, especially if the specific advantages of each domain are properly employed [39]. The same group found relevant results for the visual terminal (i.e., at the end of each task execution) feedback and visuo-haptic concurrent feedback; the auditory domain was revealed as unable to properly augment training on 3-dimensional movements [12].

F. Aim of this Work

In this work, a VR robotic teleoperated needle-driving task was integrated with different augmentation modalities both in the visual and haptic domain. Needle-driving was chosen as it both exemplifies a surgery-related set of sub-routines, and it is a complex motor learning bi-manual task that engages together strategy and visuo-motor accuracy. To optimize the augmentation of the presented task, subsequent guidance and feedback were delivered during the same training session. The rationale behind this novel approach derives from the joint consideration that guidance tends to prevent the learner from making procedural errors [40], and it should be removed before becoming integral part of the task [41]. On the other side, feedback can contribute to movement refinement through error correction [13] only if the user has previously acquired the correct procedural steps of the task. Both these augmentation modalities can be provided to the trainee in form of 3-dimensional visual cues or forces/torques delivered at the robot master manipulators.

Then, a secondary goal of this study is comparing the effectiveness of augmented guidance and feedback in different sensory domains (visual and haptics). As above mentioned, this comparison has been poorly investigated in robot-assisted surgical training in a structured way.

III. Materials and Methods

This section describes the user study that was carried out to investigate the effects of the different training augmentations. Firstly, the hardware and software employed in this work are listed, together with an overview of the VR task; an accurate description of the visual and haptic augmentations follows. Finally, the acquisition protocol and details about the evaluation metrics and their statistical analysis are reported.

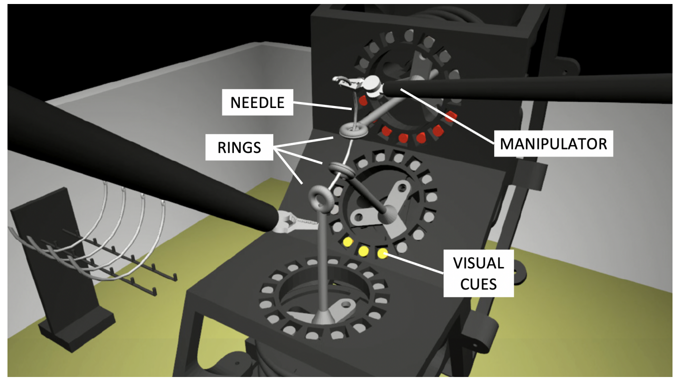


Fig. 1: The virtual END platform [36]. Three rings compose the needle insertion/extraction trajectory at the centre of the scene. In the image, the right manipulator holds the needle during insertion. In the current picture, the task is in Visual Feedback mode: the color coding and the position along each LED strip indicate the intensity and directionality of ring displacements, respectively.

A. Experimental Setup

The robotic platform employed in this study was the master console of the da Vinci Research Kit [23]. This console includes a foot-pedal tray, an HD stereo viewer and two 8DoF master manipulators, with only the first 7 joints actuated. The ATAR (Assisted Teleoperation with Augmented Reality) framework [24] was employed to design the virtual reality task, while the meshes of the virtual objects were created with SolidWorks and Blender. The simulation core runs at 60 Hz refreshing both the graphical and physical rendering engines. A separate thread running at 1kHz manages the generation of the forces and torques necessary for the haptic augmentation and, contextually, the bilateral communication with the dVRK.

B. Virtual Task

The virtual Enhanced Needle Driving (END) platform [36] was used to simulate three of the sub-routines that characterize in-vivo needle-driving: insertion, hand-to-hand transfer, and extraction. At first, a needle is grasped by the user with its dominant hand to be inserted in a series of three rings collocated at 0, 45, and 90 degrees from the horizontal plane (Fig. 1). Once correctly inserted, the needle is transferred to the non-dominant hand to be temporarily held in place. Finally, the dominant hand picks-up the needle again (close to the tip) and performs the extraction. The three rings are compliant to deviated trajectories of the needle and their displacement is used to both generate the augmented feedback and to evaluate the user's accuracy performance.

C. Training Augmentations

Visual Guidance - The visual guidance was provided by means of 3D graphic overlays opportunely positioned in the scene to guide the subject during the different task phases: needle *pick-up*, *pre-rotation*, *insertion*, *transfer*, and *extraction* (Fig. 2). Three types of overlays were implemented: a dynamic *ghost gripper*, showing the correct grasping position and orientation, as well as the

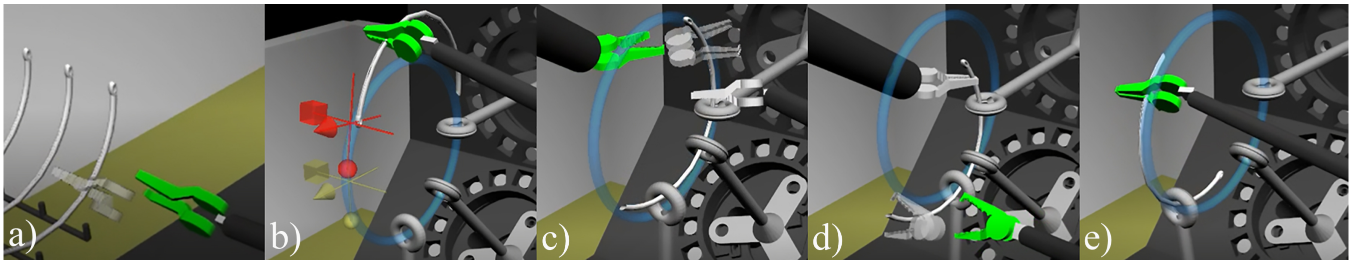


Fig. 2: Visual Guidance: a) The gray translucent gripper represents the *ghost gripper* for needle pick up; b) The red dynamic *Reference frame* indicates the current needle tail pose and has to be matched to the static yellow *Reference frame* for correct needle pre-rotation; c) The gray translucent *ghost gripper* guides the user during transfer to the non-dominant hand; d) The gray translucent *ghost gripper* is used for transfer to the dominant hand; e) The blue *torus* marks the ideal insertion and extraction trajectory. The green gripper in all the sub-images represents the current pose of the manipulator tele-operated by the user.

jaws open/close timing during *pick-up* and *transfer* (Fig. 2a,c,d); a dynamic and a static *reference frame* to be matched for univocal *pre-rotation* and positioning of the needle at the entry point (Fig. 2b); and finally, a static blue *torus* highlighting the ideal trajectory to guide the user during *insertion* and *extraction* (Fig. 2e). The visual overlays interactivity was managed through a kinematic pose matching function that checks whether two specified virtual objects (i.e. the user driven and the *ghost* grippers) correctly overlap. The task complexity can be customized by setting the matching tolerance over 6DoF; for this study the tolerance was set at 1 cm for the translational DoFs and 20° for the rotational DoFs. Reduced opacity was opportunely set for the overlays to avoid visual occlusion.

Visual Feedback - The visual feedback was developed as a multi directional real-time visual cue carrying information about the displacement of the rings. In details, at each time step the direction and the degree of deflection of the shaft supporting each ring is mapped into a circular LED strip located, co-axially, at the base of the shaft. The resulting visual feedback (Fig.1) intuitively displays intensity (color coding and number of LEDs turned on) and direction (position on the LEDs turned on) of each ring displacement. The user is therefore able to identify and correct the trajectory deviations. More details can be found at the original description of the END platform [36]. A video of the implemented visual guidance and feedback can be found in the attached media.

Haptic Guidance - For the development of the haptic guidance, a desired pose for the virtual manipulator $\boldsymbol{T_{man,d}}$ was defined at each task phase. During the *pick-up*, *transfer*, and *pre-rotation* phases, $\boldsymbol{T_{man,d}}$ was defined as static, and it corresponded to the same poses used for the visual augmentations (*ghost grippers* and *reference frame*) in Section III-C. Throughout the *insertion* and *extraction* phases a specific algorithm, similar to the work described by Coad et al. [25], was used to define $\boldsymbol{T_{man,d}}$ at each instant of time. The whole *insertion-extraction* ideal path was discretized in 300 desired needle tip poses $\boldsymbol{T_{tip,d}}$. Such discrete trajectory, as a whole, perfectly matches to the one highlighted by the visual blue *torus* mentioned in Section III-C. At first, the algorithm looks for the $\boldsymbol{T_{tip,d}}$ on the trajectory that is closest to the current needle tip pose $\boldsymbol{T_{tip,c}}$. Therefore, the chosen $\boldsymbol{T_{tip,d}}$ is multiplied for the current transformation between the needle tip and the manipulator ${}^{\boldsymbol{tip}}\boldsymbol{T_{man,c}}$ obtaining the desired $\boldsymbol{T_{man,d}}$. From here, the current and desired manipulator poses $\boldsymbol{T_{man,c,d}}$ are used to calculate the appropriate force and torque vectors to send to the dVRK masters. A simple visco-elastic method was employed for the generation of forces and torques:

$$\boldsymbol{f}^G = K_t^G * (\boldsymbol{t_{man,d}} - \boldsymbol{t_{man,c}}) - D_t^G * \boldsymbol{v_{man,c}} \quad (1)$$

$$\boldsymbol{t}^G = K_r^G * (\boldsymbol{r_{man,d}} - \boldsymbol{r_{man,c}}) - D_r^G * \boldsymbol{\omega_{man,c}} \quad (2)$$

where $\boldsymbol{t_{man,d,c}}$ are vectors of x, y, z coordinates and $\boldsymbol{r_{man,d,c}}$ are vectors of $roll, pitch, yaw$ angles extracted from $\boldsymbol{T_{man,d,c}}$. Furthermore, $\boldsymbol{v_{man,c}}$ and $\boldsymbol{\omega_{man,c}}$ are the manipulator current velocity and angular rate, while $K_{t,r}^G$ and $D_{t,r}^G$ are translational and rotational elastic and viscosity coefficients, respectively. The coefficient values, as well as the maximum applied forces and torques are reported in Table I. To avoid jumps or system instabilities, a ramp function ($\in 0, 1$) was applied to wrenches during the first Δt (1 s) of guidance activation in each phase. Since the opening/closure of the gripper is not actuated in the dVRK platform, no haptic guidance nor feedback was provided to the trainee in terms of grasping.

TABLE I: Haptic guidance parameters

K_t^G	K_r^G	D_t^G	D_r^G	f_{max}^G	t_{max}^G
180 N/m	0.15 N/m	15 N.s/m	0.01 N.s/m	4 N	0.1 N.m

Haptic Feedback - A visco-elastic method was also used for the haptic feedback. At any time instant, the reference pose (at equilibrium) $\boldsymbol{T_{ring,r}}$ and current pose (when displaced) $\boldsymbol{T_{ring,c}}$ of each ring were compared; as a result, three distance vectors $\boldsymbol{d_i}$, one for each ring, were computed. The aim of the haptic feedback is that of delivering to the user a clear and realistic feeling of the interaction with the environment (in this case the needle-rings interactions). Therefore, considering the complexity of estimating accurately the cumulative torques resulting at the manipulator from each directional component ($d_{i,x,y,z}$ of each $\boldsymbol{d_i}$ applied with a different torque arm) and the marginal gain that such computation could have introduced [42], the model was simplified to forces-only:

$$\boldsymbol{f}^F = \sum_{i=1}^{3} \boldsymbol{d_i} * K_t^F - D_t^F * \boldsymbol{v_{man,c}} \quad (3)$$

where *i* spans the number of rings. The visco-elastic parameters used for the generation of the force vector **f**, as well as the maximum applied values are reported in Table II.

TABLE II: Haptic feedback parameters

K_t^F	D_t^F	f_{max}^F
360 N/m	40 N.s/m	3 N

D. Acquisition protocol

The user study involved 32 non-medical participants (from 21 to 35 years old, 19 males and 13 females). All the subjects had none to little experience with teleoperated robots, where little experience referred to the use of a tele-operated robot just during demonstrations and not during past user studies. They were randomly divided into 4 groups: a control group (C) which did not receive any kind of guidance or feedback, a visual group (V) which performed training with visual guidance and feedback, a haptic group (H) which underwent haptic guidance and feedback, and a visuo-haptic group (VH) which had cues of both the sensory domains.

Each experimental session was comprised of three phases: a baseline assessment (B, composed by 5 task repetitions), a training stage (G and F, 10 repetitions each) and a final evaluation (E, 5 repetitions). All the groups performed the task with no guidance/feedback during baseline and evaluation. Only the training stage differed among groups. The control users had no cue. The training phase of the other groups was split into two sub-stages: in the former (G, 10 repetitions), each group experienced guidance of the respective sensory domain (visual, haptic, or both); in the latter (F, 10 repetitions), they received feedback of the assigned sensory domain.

At the beginning of the experiment, the subjects were introduced to the dVRK console. Each user (independently from the allocation group) was shown a video with a correct execution of the task and received a verbal explanation of its salient steps. Moreover, all participants were made aware that their performance would be evaluated in terms of ring displacement minimization. The experiments were approved by the ethical committee of Politecnico di Milano, and all the subjects gave informed consent according to the declaration of Helsinki.

E. Performance metrics and statistical analysis

The first metric, *Task Completion*, related to the capability of the user to complete each task phase (needle insertion, transfer, and extraction). For each task repetition, a binary index stating the conclusion of each task phase was recorded. The *Task Completion* of a certain user in a certain task repetition is the sum of completed phases out of the total number of phases (i.e., 3).

Secondly, the time-accuracy performance was addressed by computing the *Time* to complete the task and the *Mean Ring Displacement* during the task execution. The *Mean Ring Displacement* is defined as the average of the sum of the euclidean distances of the 3 rings from their resting positions during the task [36].

Whenever the participant was not able to complete the full task, the trial was considered as incomplete and *Mean Ring Displacement* and *Time* data were ignored. A total of 334 trials, corresponding to the 35% of the overall time-accuracy data were discarded.

Considering this incomplete nature of the analyzed data set, a Generalized Linear Mixed Model (GLMM) was chosen as the most suitable fit for the repeated measures design of the experiment. The model was set to take into account for the interaction of the two independent variables (group and stage) while subjects were treated as a random effect. The three performance metrics evaluated in this work were alternatively set in the GLMM as dependent variable. Given that the resulting distribution of the dependent variables was not normally distributed, a log10 transformation was applied to each *Time* and *Mean Ring Displacement* data point to restore normality of the models residuals. Data fitting was done using Maximum Likelihood Estimation (MLE). Multiple linear hypothesis testing was performed on the GLMM using simultaneous t-tests with Satterthwaite's method; Bonferroni corrections were applied to account for multiple comparisons in the data sets. All statistical analyses were performed in R version 3.5.3.

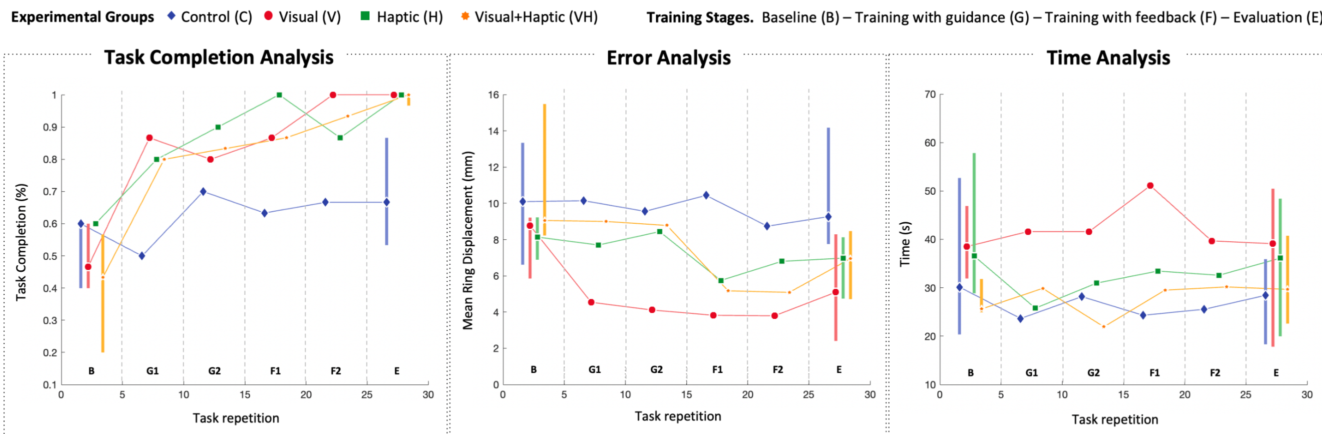
IV. Results and Discussion

This section reports the main outcomes in terms of effectiveness of training augmentation (guidance and feedback), as well as a comparison of sensory domains. This effectiveness is addressed in terms of capability to complete the task (*Task Completion*), as well as time-accuracy metrics (*Mean Ring Displacement*) and *Time*). All the subjects enrolled in the user study completed the acquisition protocol. The duration of a single session ranged from 46 up to 72 minutes.

A. Task Completion

Despite the quality of execution, the basic objective of learning a task is at least being able to complete it. In the VR exercise under investigation, needle drop is the key source of failure. Such a failure is likely to occur when the needle is grasped or transferred from instrument to instrument in sub-optimal ways. These procedural aspects can be complex to take into account for novice trainees.

The introduction of guidance at the very beginning of training has the primary goal of assisting the trainee throughout the task procedural steps. The *Task Completion Analysis* left box of Fig. 3 shows the percentage of task completion of each group across the different training stages. The baseline assessment (B) featured similar initial values among the different experimental groups ($p>0.05$). When training



Task Completion Analysis

Statistical Analysis Within Groups

p	B vs G1	G1 vs G2	G2 vs F1	F1 vs F2	F2 vs E	B vs E
C	N	N	N	N	N	N
V	***	N	N	N	N	***
H	**	N	N	N	N	***
VH	***	N	N	N	N	***

Statistical Analysis Between Groups

p	B	G1	G2	F1	F2	E
C vs V	N	**	N	N	*	***
C vs H	N	**	*	***	*	***
C vs VH	N	*	*	*	*	***
V vs H	N	N	N	N	N	N
V vs VH	N	N	N	N	N	N
H vs VH	N	N	N	N	N	N

Error Analysis

Statistical Analysis Within Groups

p	B vs G1	G1 vs G2	G2 vs F1	F1 vs F2	F2 vs E	B vs E
C	N	N	N	N	N	N
V	*	N	N	N	N	***
H	N	N	***	N	N	***
VH	N	N	***	N	*	***

Statistical Analysis Between Groups

p	B	G1	G2	F1	F2	E
C vs V	N	N	*	***	*	**
C vs H	N	N	N	N	N	*
C vs VH	N	N	N	*	N	*
V vs H	N	N	*	N	N	N
V vs VH	N	N	*	N	N	N
H vs VH	N	N	N	N	N	N

Time Analysis

Statistical Analysis Within Groups

p	B vs G1	G1 vs G2	G2 vs F1	F1 vs F2	F2 vs E	B vs E
C	N	N	N	N	N	N
V	N	N	*	**	N	N
H	**	*	***	N	N	N
VH	N	***	***	N	N	N

Statistical Analysis Between Groups

p	B	G1	G2	F1	F2	E
C vs V	N	N	N	N	N	N
C vs H	N	N	N	N	N	N
C vs VH	N	N	N	N	N	N
V vs H	N	N	N	N	N	N
V vs VH	N	N	N	N	N	N
H vs VH	N	N	N	N	N	N

Statistical measures N = no statistical difference; * = p<0.05; ** = p < 0.01; *** = p < 0.001

Fig. 3: Graphical and statistical analysis of the *Task Completion* (left box), *Mean Ring Displacement* (central box) and *Time* (right box) metrics. On the top, each vertical box features the graphical evolution of the metrics. Markers refer to the median across the users composing each group, while the vertical bars in baseline and evaluation stand for data variability (25th and 75th percentiles). For each training slot *Task Completion* was averaged across the 5 repetitions for each user, while all the available values of the *Mean Ring Displacement* and *Time* were taken into account. On the bottom, each box shows the statistical analysis within groups across training stages (upper table) and the statistical analysis across groups for each training stage (lower table).

started (G1), the capability of task completion featured a steep jump in the groups undergoing augmented guidance with respect to the control group. This initial improvement in *Task Completion* was statistically significant across all the experimental groups ($p=0.002$ for H and $p<0.001$ for V and VH) and it resulted in a significant difference in G1 between the control users and all the experimental subjects ($p=0.004$, $p=0.006$, $p=0.041$ for V, H, and VH respectively).

As the training session continued, no further significant improvement in *Task Completion* was denoted ($p>0.05$ for all the groups and the training stages after G1). Specifically, the application of feedback instead of guidance during F1 and F2 did not seem to have significant effects on *Task Completion*. At the end of training (E), all the experimental groups outperformed the control one ($p<0.001$) and no significant difference was present across the experimental groups ($p>0.05$) at evaluation.

To summarize, the initial introduction of guidance caused a steep improvement in *Task Completion* for the experimental users. After this training stage (G1), the control users did not manage to fill the gap with respect to the other experimental groups. The sensory domain of the augmentations did not have any influence in terms of *Task Completion* capabilities.

B. *Mean Ring Displacement*

Especially in delicate fields like surgery, accuracy has a remarkable relevance. The *Error Analysis* central box of Fig. 3 depicts the evolution of *Mean Ring Displacement* across the experimental session. All the groups featured similar starting performance ($p>0.5$) in terms of accuracy at baseline (B). From baseline (B) to evaluation (E), just the experimental groups significantly improved their performance ($p<0.001$). This error reduction resulted in significantly lower values in the final *Mean Ring Displacement* of the experimental subjects with respect to the control users ($p=0.003$, $p=0.036$, $p=0.037$). Comparing the experimental groups at evaluation (E), the *Mean Ring Displacement* of the visual group was lower with respect to the haptics and visuo-haptics groups but not in a statistically significant way ($p>0.05$). The analysis of *Mean Ring Displacement* evolution during training might provide additional insights to better understand these final outcomes.

The initial research hypothesis of this study is that the application of feedback is able to minimize the error while performing the task. This was verified for the haptics and visuo-haptics subjects who underwent a significant improvement at the introduction of feedback ($p<0.001$ at G2vsF1).

Nevertheless, the visual group featured a different trend and it showed the major improvement at the introduction of guidance ($p=0.034$ at BvsG1). This could be explained by the fact that the visual guidance augmentations intrinsically contained accuracy information by showing the ideal trajectory (for the needle insertion and extraction) that causes no ring displacement.

Additionally, the lack of improvement during feedback augmentation could be a consequence of the way visual feedback was provided to the users: each ring was surrounded by LED strips whose color coding and position signaled intensity and directionality of ring displacements, respectively. This way to provide feedback could be too complex for the users to interpret, thus causing no further improvement in accuracy.

Finally, the comparison between the final stage of feedback (F2) and the evaluation (E) could give some hints about the eventual addiction of the users to the augmentations: although the median error was higher in E with respect to F2, such a deterioration was significant just for the visual-haptics group ($p=0.017$ at F2vsE). This could be explained by the simultaneous removal of a double source of augmentation in the subjects previously undergoing both visual and haptic feedback.

C. Time

The total time to complete the task was not a metric the users were verbally thought to optimize, neither they were provided feedback about. Anyway, its analysis could highlight eventual relations between time and augmentations. This analysis is reported in the *Time Analysis* right box of Fig. 3. Firstly, all the groups showed similar performance ($p>0.05$) both at baseline (B) and evaluation (E). At the same time, no significant difference between B and E was found in any group. This could lead to the conclusion that subjects did not optimize the time to complete the task throughout the training, as well as the augmentations did not help in achieving any time improvement. The former consideration can be related to the time-accuracy trade off: as previously described while discussing the *Mean Ring Displacement*, users improved their accuracy throughout the training and this could have required higher time to complete the task. As a consequence, any time improvement in terms of users' confidence in carrying out the task could be counterbalanced by additional time spent for error minimization at the end of the experiment.

Anyway, some interesting considerations could be derived by analysing the *Time* evolution within groups during training. All the experimental subjects were characterized by a relevant deterioration at the introduction of feedback ($p=0.042$ for V and $p<0.001$ for H and VH at G2vsF1). This could be related to the cognitive demand in initially interpreting the feedback. Focusing on the haptics group, it was subjected to a significant *Time* improvement at the beginning of guidance ($p=0.009$ at BvsG1). This could be linked to the application of forces that can speed up the task execution by triggering movements and accelerating motion.

V. Conclusions

Virtual reality simulation and robotics (as visual and haptic interfaces, respectively) open attractive opportunities for training augmentation. This work focused on the integration of a virtual reality surgical task (needle driving) with different augmentation modalities to optimize training. Such augmentations were provided to the user as a sequential combination of guidance (showing how to carry out the task) and feedback (showing the concurrent error), and they were implemented in the visual and haptic domain. A control group (performing training without any augmentation) was compared to three experimental groups (receiving training augmentation in the visual, haptic and both the domains, respectively).

All the subjects undergoing augmentations outperformed the control group at the end of training in terms of task completion and accuracy. More specifically, the presence of guidance at the beginning of training played as determining factor for all the experimental groups in terms of task completion, and additionally as a source of error reduction when deployed in the visual domain. The subsequent application of feedback did not cause a comparable further improvement in task completion for any of the groups, while haptic and visuo-haptic groups significantly reduced the error when the concurrent feedback was applied. These results hint at the potential for future experimentation in coupling visual guidance and haptic feedback during subsequent training stages to further optimize learning curves during RAMIS training. No significant reduction was shown in terms of total task execution time from pre to post training assessments, .

Despite demonstrating a distinct aid of training augmentation, few significant differences were identified when comparing sensory domains. This limitation could be related to the short duration of training (single session), as well as the lack of skill retention tests.

Further investigations will move in this direction by defining a longer study protocol, which includes multiple sessions over different days. To address this purpose, the complexity of the target task to learn should be increased (in order to prevent early saturation of learning). This opens the possibility to deal with more structured task of robotic surgery, involving challenging cognitive loads due to decision making and emergency response. The contribution of augmentation in such situations could be an additional interesting insight to address both from a sensory-motor and a psychological point of view.

Additionally, shifting research from single gestures of surgery (like needle driving) to full steps of surgical procedures can pave the way towards a full exploitation of simulation-based training even after the very initial practical skill development as in the current practise. Such an ideal training platform enriched with multi-sensorial automated coaching for guidance and feedback could optimize training outcomes, decrease learning times, as well as cut costs associated to mentoring.

ACKNOLEDGMENT

The authors would like to thank Dr. Jeremy D. Brown and Dr. Mahya Shahbazi for the early development of this work; Eric Cao, Vipul Bhat, Brett Wolfinger, Anna Bucchieri, Giulia Bruni and Aicha Bourkadi Idrissi for their contributions.

REFERENCES

[1] S. Maeso, M. Reza, J. A. Mayol, J. A. Blasco, M. Guerra, E. Andradas, and M. N. Plana, "Efficacy of the Da Vinci surgical system in abdominal surgery compared with that of laparoscopy: a systematic review and meta-analysis," 2010.

[2] 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, 2014.

[3] J. E. Thompson, S. Egger, M. Böhm, A.-M. Haynes, J. Matthews, K. Rasiah, and P. D. Stricker, "Superior quality of life and improved surgical margins are achievable with robotic radical prostatectomy after a long learning curve: a prospective single-surgeon study of 1552 consecutive cases," *Euro. uro.*, vol. 65, no. 3, pp. 521–531, 2014.

[4] A. Aydin, K. Ahmed, A. M. A. Shafi, M. S. Khan, and P. Dasgupta, "The role of simulation in urological training – A quantitative study of practice and opinions," *The S.*, vol. 14, no. 6, pp. 301–307, 2016.

[5] F. Schlottmann, S. Tolleson-Rinehart, M. R. Kibbe, and M. G. Patti, "Status of Simulation-Based Training in Departments of Surgery in the United States," *J. of surgical research*, vol. 255, pp. 158–163, 2020.

[6] F. H. Halvorsen, O. J. Elle, V. V. Dalinin, B. E. Mørk, V. Sørhus, J. S. Røtnes, and E. Fosse, "Virtual reality simulator training equals mechanical robotic training in improving robot-assisted basic suturing skills," *S. End. and O. Int. Tec.*, vol. 20, no. 10, pp. 1565–1569, 2006.

[7] D. Escobar-Castillejos, J. Noguez, F. Bello, L. Neri, A. J. Magana, and B. Benes, "A Review of Training and Guidance Systems in Medical Surgery," *Applied Sciences*, vol. 10, no. 17, p. 5752, 2020.

[8] F. G. Hamza-Lup, C. M. Bogdan, D. M. Popovici, and O. D. Costea, "A survey of visuo-haptic simulation in surgical training," *prepr.*, 2019.

[9] M. Connolly, J. Seligman, A. Kastenmeier, M. Goldblatt, and J. C. Gould, "Validation of a virtual reality-based robotic surgical skills curriculum," *Surgical endoscopy*, vol. 28, no. 5, pp. 1691–1694, 2014.

[10] R. Hatala, D. A. Cook, B. Zendejas, S. J. Hamstra, and R. Brydges, "Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis," *Advances in Health Sciences Education*, vol. 19, no. 2, pp. 251–272, 2014.

[11] A. W. Salmoni, R. A. Schmidt, and C. B. Walter, "Knowledge of results and motor learning: a review and critical reappraisal." *Psychological bulletin*, vol. 95, no. 3, p. 355, 1984.

[12] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Terminal feedback outperforms concurrent visual, auditory, and haptic feedback in learning a complex rowing-type task," *Journal of motor behavior*, vol. 45, no. 6, pp. 455–472, 2013.

[13] C. K. Williams and H. Carnahan, "Motor learning perspectives on haptic training for the upper extremities," *IEEE trans. on haptics*, vol. 7, no. 2, pp. 240–250, 2014.

[14] B. Zhu, D. Kaber, M. Zahabi, and W. Ma, "Effects of feedback type and modality on motor skill learning and retention," *Behaviour & Information Technology*, vol. 39, no. 4, pp. 431–442, 2020.

[15] J. A. Saunders and D. C. Knill, "Visual feedback control of hand movements," *J. of Neuroscience*, vol. 24, no. 13, pp. 3223–3234, 2004.

[16] C. E. Reiley, T. Akinbiyi, D. Burschka, D. C. Chang, A. M. Okamura, and D. D. Yuh, "Effects of visual force feedback on robot-assisted surgical task performance," *The J. of thoracic and cardiovascular surgery*, vol. 135, no. 1, pp. 196–202, 2008.

[17] T. N. Judkins, D. Oleynikov, and N. Stergiou, "Real-time augmented feedback benefits robotic laparoscopic training," *Studies in health technology and informatics*, vol. 119, p. 243, 2005.

[18] "Mimic Technologies, Inc., WA 98104, USA." [Online]. Available: https://mimicsimulation.com/da-vinci-skills-simulator/

[19] A. Malpani, S. S. Vedula, H. C. Lin, G. D. Hager, and R. H. Taylor, "Real-time teaching cues for automated surgical coaching," 2017.

[20] K. Rangarajan, H. Davis, and P. H. Pucher, "Systematic Review of Virtual Haptics in Surgical Simulation: A Valid Educational Tool?" *J. of surgical education*, vol. 77, no. 2, pp. 337–347, 2020.

[21] D. Pinzon, S. Byrns, and B. Zheng, "Prevailing trends in haptic feedback simulation for minimally invasive surgery," *Surgical innovation*, vol. 23, no. 4, pp. 415–421, 2016.

[22] M. Shahbazi, S. F. Atashzar, C. Ward, H. A. Talebi, and R. V. Patel, "Multimodal sensorimotor integration for expert-in-the-loop telerobotic surgical training," *IEEE Transactions on Robotics*, vol. 34, no. 6, pp. 1549–1564, 2018.

[23] 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 ICRA*. IEEE, 2014, pp. 6434–6439.

[24] N. Enayati, A. M. Okamura, A. Mariani, E. Pellegrini, M. M. Coad, G. Ferrigno, and E. De Momi, "Robotic assistance-as-needed for enhanced visuomotor learning in surgical robotics training: An experimental study," in *2018 IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 6631–6636.

[25] M. M. Coad, A. M. Okamura, S. Wren, Y. Mintz, T. S. Lendvay, A. M. Jarc, and I. Nisky, "Training in divergent and convergent force fields during 6-DOF teleoperation with a robot-assisted surgical system," in *2017 IEEE World Haptics Conf. (WHC)*. IEEE, 2017, pp. 195–200.

[26] H. Heuer and J. Lüttgen, "Robot assistance of motor learning: A neuro-cognitive perspective," *Neu. & Biob. Rev.*, vol. 56, pp. 222–240, 2015.

[27] P. Schleer, P. Kaiser, S. Drobinsky, and K. Radermacher, "Augmentation of haptic feedback for teleoperated robotic surgery," *Int. J. of Comp. Assisted Rad. and Surg.*, vol. 15, no. 3, pp. 515–529, 2020.

[28] J. D. Brown, J. N. Fernandez, S. P. Cohen, and K. J. Kuchenbecker, "A wrist-squeezing force-feedback system for robotic surgery training," in *2017 IEEE World Haptics Conf. (WHC)*. IEEE, 2017, pp. 107–112.

[29] K. J. Kuchenbecker, J. Gewirtz, W. McMahan, D. Standish, P. Martin, J. Bohren, P. J. Mendoza, and D. I. Lee, "VerroTouch: High-Frequency Acceleration Feedback for Telerobotic Surgery," in *Haptics: Generating and Perceiving Tangible Sensations*, A. M. L. Kappers, J. B. F. van Erp, W. M. Bergmann Tiest, and F. C. T. van der Helm, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 189–196.

[30] J. Franks, M. Culjat, C.-H. King, M. Franco, J. Bisley, W. Grundfest, and E. Dutson, "Pneumatic balloon actuators for tactile feedback in robotic surgery," *Industrial Robot: An Int. J.*, 2008.

[31] A. Abiri, J. Pensa, A. Tao, J. Ma, Y.-Y. Juo, S. J. Askari, J. Bisley, J. Rosen, E. P. Dutson, and W. S. Grundfest, "Multi-modal haptic feedback for grip force reduction in robotic surgery," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[32] Z. Chen, A. Malpani, P. Chalasani, A. Deguet, S. S. Vedula, P. Kazanzides, and R. H. Taylor, "Virtual fixture assistance for needle passing and knot tying," in *2016 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 2343–2350.

[33] A. Banach, K. Leibrandt, M. Grammatikopoulou, and G.-Z. Yang, "Active contraints for tool-shaft collision avoidance in minimally invasive surgery," in *2019 IEEE ICRA)*. IEEE, 2019, pp. 1556–1562.

[34] N. Schaffert, T. B. Janzen, K. Mattes, and M. H. Thaut, "A review on the relationship between sound and movement in sports and rehabilitation," *Frontiers in psychology*, vol. 10, p. 244, 2019.

[35] 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, pp. 2970–2983, 2015.

[36] G. Caccianiga, A. Mariani, E. De Momi, G. Cantarero, and J. D. Brown, "An Evaluation of Inanimate and Virtual Reality Training for Psychomotor Skill Development in Robot-Assisted Minimally Invasive Surgery," *IEEE T-MRB*, vol. 2, no. 2, pp. 118–129, 2020.

[37] H. U. Yoon, R. F. Wang, S. A. Hutchinson, and P. Hur, "Customizing haptic and visual feedback for assistive human–robot interface and the effects on performance improvement," *Robotics and Autonomous Systems*, vol. 91, pp. 258–269, 2017.

[38] L. Marchal-Crespo, M. van Raai, G. Rauter, P. Wolf, and R. Riener, "The effect of haptic guidance and visual feedback on learning a complex tennis task," *Exp. br. res.*, vol. 231, no. 3, pp. 277–291, 2013.

[39] R. Sigrist, G. Rauter, L. Marchal-Crespo, R. Riener, and P. Wolf, "Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning," *Experimental brain research*, vol. 233, no. 3, pp. 909–925, 2015.

[40] D. H. Holding and A. W. Macrae, "Guidance, restriction and knowledge of results," *Ergonomics*, vol. 7, no. 3, pp. 289–295, 1964.

[41] J. C. Huegel and M. K. O'Malley, "Progressive haptic and visual guidance for training in a virtual dynamic task," in *2010 IEEE Haptics Symposium*. IEEE, 2010, pp. 343–350.

[42] L. Santos-Carreras, R. Beira, A. Sengül, 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, 2010.