

RESEARCH ARTICLE

Open Access



A haptic laparoscopic trainer based on affine velocity analysis: engineering and preliminary results

Benjamin De Witte^{1†}, Charles Barnouin^{2†}, Richard Moreau², Arnaud Lelevé^{2*} , Xavier Martin^{3,4}, Christian Collet¹ and Nady Hoyek¹

Abstract

Background: There is a general agreement upon the importance of acquiring laparoscopic skills outside the operation room through simulation-based training. However, high-fidelity simulators are cost-prohibitive and elicit a high cognitive load, while low-fidelity simulators lack effective feedback. This paper describes a low-fidelity simulator bridging the existing gaps with affine velocity as a new assessment variable. Primary validation results are also presented.

Methods: Psycho-motor skills and engineering key features have been considered e.g. haptic feedback and complementary assessment variables. Seventy-seven participants tested the simulator (17 expert surgeons, 12 intermediates, 28 inexperienced interns, and 20 novices). The content validity was tested with a 10-point Likert scale and the discriminative power by comparing the four groups' performance over two sessions.

Results: Participants rated the simulator positively, from 7.25 to 7.72 out of 10 (mean, 7.57). Experts and intermediates performed faster with fewer errors (collisions) than inexperienced interns and novices. The affine velocity brought additional differentiations, especially between interns and novices.

Conclusion: This affordable haptic simulator makes it possible to learn and train laparoscopic techniques. Self-assessment of basic skills was easily performed with slight additional cost compared to low-fidelity simulators. It could be a good trade-off among the products currently used for surgeons' training.

Keywords: Minimal invasive surgery, Motor skills, Assessment metrics, Affine velocity, Simulator validation

Background

Minimally Invasive Surgery (MIS) has brought more comfort for the patients in comparison to open surgery [1–3]. However, MIS also challenged surgeons with an increased need for manual dexterity, depth perception, and movement coordination through a 2D screen displaying the surgical field. Therefore, efficient training

of spatial abilities [4], bi-manual coordination [5], and hand–eye coordination [6], are essential for young surgeons [7, 8]. During the past two decades, simulators have been more extensively used in surgeons' training [9], without the risk of harming patients [10, 11], with health care cost reductions [10], and better efficiency than with simple observation [12]. Two kinds of simulators coexist: low and high fidelity simulators. Low-fidelity ones use simple real instruments and are better suited to novice learners [11] as they are efficient for basic skills acquisition [13] such as clipping, grasping, or cutting. However, they lack fidelity [14] and fail to provide immediate and summative feedback [14–16], which seriously hinders

*Correspondence: arnaud.leleve@insa-lyon.fr

[†]Benjamin De Witte and Charles Barnouin contributed equally to this study

² INSA Lyon, Ampère (UMR5005), Univ Lyon, 25 av. Jean Capelle ouest, 69621 Villeurbanne Cedex, France

Full list of author information is available at the end of the article



learning [11, 17, 18]. Also, they require a subjective observation by an expert [14, 19], which reduces effective training opportunities for novice surgical students [20]. High-fidelity simulators, such as the LAP Mentor™ (Simbionix Corporation, Cleveland, Ohio, USA) improve the performance of full complex laparoscopic procedures (such as hysterectomy [21]) [9]. Nevertheless, they are very cost-prohibitive and may not be accessible for regular and personal use [19, 22, 23].

Even if low- and high-fidelity simulators may be considered as a continuum in the learning process [13, 17], a gap remains between them: for instance, Yiasemidou et al. [24] showed that the students who autonomously trained on a “take-home” box trainer (\approx \$500) at home during 6 weeks performed better than the ones who practiced on the High-Fidelity Virtual-Reality-based simulator (\approx \$70,000). Therefore, a solution mixing low-cost, broad availability, autonomous use, and objective automatic assessment should efficiently fill this gap.

To address the above-mentioned flaws, this study aims to develop and validate a new low-cost (\approx \$2500) low-fidelity simulator providing: (i) training on basic laparoscopic psycho-motor tasks (requiring spatial and visual-motor capabilities to acquire universal gesture-based skills versus specific surgical ones such as in [25]), (ii) both haptic and summative feedback, and (iii) a new assessment variable (in a similar approach as in [26]) permitting the evaluation of the smoothness of motions and so better expertise discrimination. As shown by Shout et al. [27], before testing the training efficiency of a simulator, scientific validation needs to be processed. According to current standards for educational and psychological testing [28, 29], we present the preliminary validation results (reliability, content, and relation with experimental designs).

Methods

Engineering and exercises

This simulator was designed by observing surgeons in operation rooms and isolating key-skills with the help of the Fundamentals of Laparoscopic Surgery [30]. Prono-supination, elbow flexion and extension, wrist rotations, and index finger rotations were the basic anatomical movements that were analyzed. The exercises embedded into this simulator require the user to regularly perform these anatomical movements as they consist of navigation tasks using laparoscopic grasping tools while avoiding 3D obstacles displayed on a computer screen. In total, five obstacles are displayed along the trajectory, each one requiring the execution of a specific action (Fig. 1). After the design, the exercises have been submitted to and validated by the Head of Lyon Surgery Department. They combine memory work and skills involving

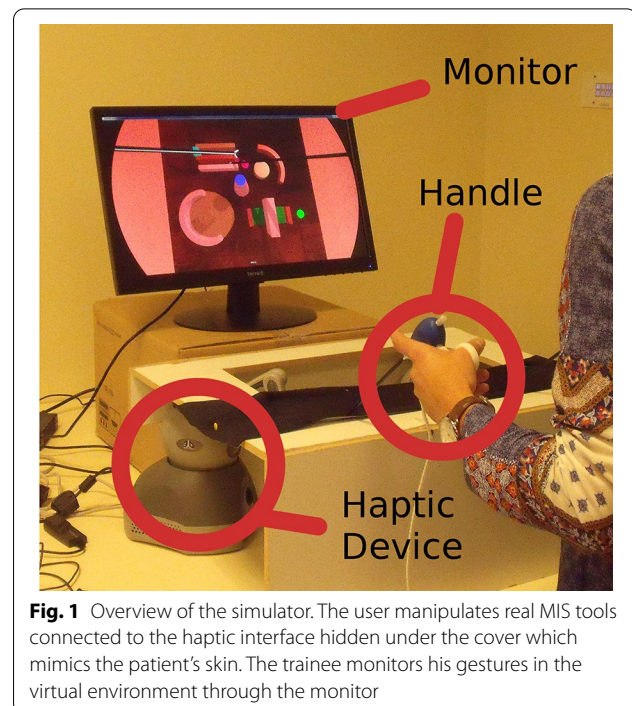


Fig. 1 Overview of the simulator. The user manipulates real MIS tools connected to the haptic interface hidden under the cover which mimics the patient's skin. The trainee monitors his gestures in the virtual environment through the monitor

visual-spatial ability [31]. Considering the high cognitive load involved by high-fidelity simulators, which may negatively impact the ability to learn medical gestures [32], we followed the recommendations of Mayer and Moreno [33] and kept the exercises and the virtual environment simple, focusing on the gestures (trajectories and tool synchronization).

To avoid high computing costs, we used open-source software and affordable material [34]. This low-fidelity simulator is a computer-based training system. The hardware includes a computer (Microsoft Windows 7® Intel® Core™ i7-6500 CPU clocked at 2.50GHz, Nvidia® Geforce GTX940M, 8Go RAM), and a standard monitor displaying virtual 3D environment, associated with two real standard laparoscopic surgical instruments, paired each one with a force-feedback device thus simulating tactile and kinesthetic feedback (Geomagic Touch® haptic device, 3D System Inc.). The latter is a 6 degrees-of-freedom device using three electrical motors as actuators, often used in medical simulators [35]. We developed the software with Microsoft Visual Studio® 2015¹ and we used the open-source Haptic Framework CHAI3D 3.1.1² to make the haptic devices, the 3D virtual world, and the Open-GL 4.4³ renderer communicating with each other.

¹ see <https://www.microsoft.com/france/visual-studio/>.

² see <http://www.chai3d.org/>.

³ see <https://www.opengl.org/>.

We designed all virtual objects with the 3DS Max software 2016.⁴ We finally collected and processed performance data (affine velocity, motion duration, number of collisions) through Matlab[®] R2015a (Mathworks).⁵

Assessment metrics used in surgical training simulators: completion time, instrument path length, number of collisions (between the laparoscopic tools and their environment), or count of dropped objects, are indirect indicators of laparoscopic proficiency that need to be associated with other objective measures i.e. motion characteristics [15, 36]. A review of existing metrics [26] concludes that available metrics do not allow an objective determination of the detailed expertise of subjects. Moreover, users do not consider these metrics consistent with their performance [37]. Indeed, experiments revealed that experts perform smoother, more accurate, and fine movements compared to novices [38], but usual metrics do not directly evaluate the curvature of instrument motions.

Affine velocity is a metric that takes the geometry of the trajectory into account, as well as its dynamics. In the 2D drawing, humans tend to decrease the instantaneous tangential velocity of their hands while the curvature of the trajectory increases. Correspondingly, the hand velocity increases when the trajectory becomes straight [39]. Furthermore, this relationship conforms to an empirical two-thirds power law [40]. For an MIS-tool 3D trajectory, this property had to be adapted for spatial motion. Some experimental results suggested that the two-thirds power law does not fit 3D motions and that a one-sixth power law is needed [41]. The relation between the Euclidean velocity v , the curvature κ , and the torsion τ is defined as:

$$v = v_a \cdot \kappa \cdot \alpha \cdot |\tau| \cdot \beta \quad (1)$$

where v_a is the affine velocity, α and β are two parameters to be determined according to the skill to be studied. For example, in [42], the motion of an obstetrical gesture (typically the installation of forceps during childbirth) was very different from those in MIS. As a result, the parameters used in these previous works could not be used for MIS. Therefore, we had to determine those for MIS, which should fit every motion involved in this hands-on training. This is explained in the next subsection. Affine velocity was utilized to provide a quantitative measure reflecting the quality of the kinematics of users' trajectories. Affine velocity was used instead of the velocity itself because it provides valuable information about the quality and smoothness of the trajectory. Data from

each training session (trajectories, moments, and numbers of collisions, right/left-hand distributions) were collected through CHAI3D into a file and post-processed using Matlab to evaluate the trainees' performance. In future versions, these assessment algorithms will be integrated into the simulator software.

The first step to tune the affine velocity computation was to get enough data to determine the best α and β values for this simulator. We collected an adequate sample of 46 trajectories ranging from novices to experts (see the next section). We independently verified their skill level so that they could serve as references. After the process proposed in [42], we first interpolated the trajectories of both tooltips into cubic splines and computed the values of the Euclidean velocity v , the curvature κ , and the torsion τ (measuring how much the trajectory twists out of the plane of curvature). Then, we performed a logarithmic linearization of (1):

$$\log(v) = \log(v_a) + \alpha \cdot \log(\kappa) + \beta \cdot \log(|\tau|) \quad (2)$$

We performed multivariate linear regression using a gradient descent algorithm to determine these parameters. The average values of α and β should fit the best laparoscopic medical skills to allow a realistic computation of affine velocity. For this medical skill, we found $\alpha = -0.048$ and $\beta = -0.0026$.

Validation processes

According to the American Psychological Association guidelines [43] and standards [28], we validated the simulator's content evidence and discriminative power.

Content evidence validation and its internal consistency

We examined if the simulator's items completely represent a basic surgical skill-learning tool. For the first exercise (Fig. 2), participants should replicate the trajectory as fast as possible with their dominant hand only, the non-dominant hand staying motionless while holding the second laparoscopic tool. During the second exercise (Fig. 3), participants had to perform the same trajectory as fast as possible with their dominant hand while the non-dominant hand should simultaneously complete a small trajectory. After completing the two exercises, the participants were invited to globally assess the simulator using a 25-item questionnaire on a 10-point rating scale. Six questions evaluated the general aspect (such as "How would you rate the height of the simulator regarding a real intervention?"), five questions were about the didactic components (How would you rate the quality of the feedback (metrics, score...)?), four questions were about the perception of reliability (How would you rate the accuracy of the movements regarding reality?), seven questions were linked to the training and learning power (Do you

⁴ see <http://www.autodesk.fr/products/3ds-max/overview>.

⁵ see <https://fr.mathworks.com>.

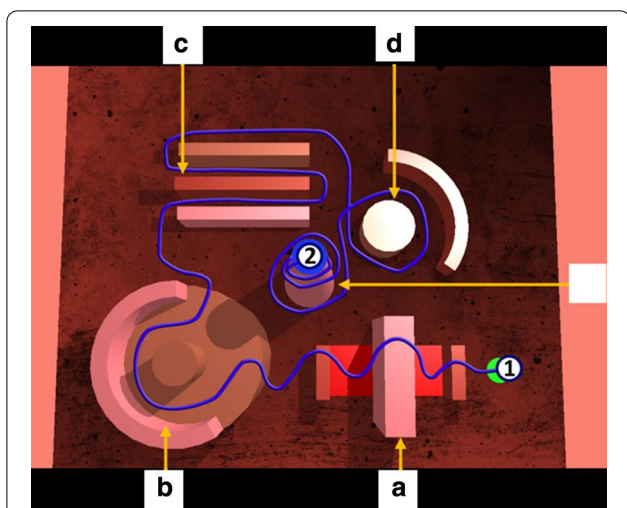


Fig. 2 Screenshot of the trajectory to perform in exercise 1 from the starting position (1) to the ending position (2). The trajectory had to be performed without touching the 3D structures. The five obstacle avoidance requested implicitly the following skills: grasping and releasing the green sphere to start, jumping over the central block in A (insertion, withdrawal), taking a curve in B, moving straight between the blocks in C, going round in D, and E, simulating loops necessary for knot tying

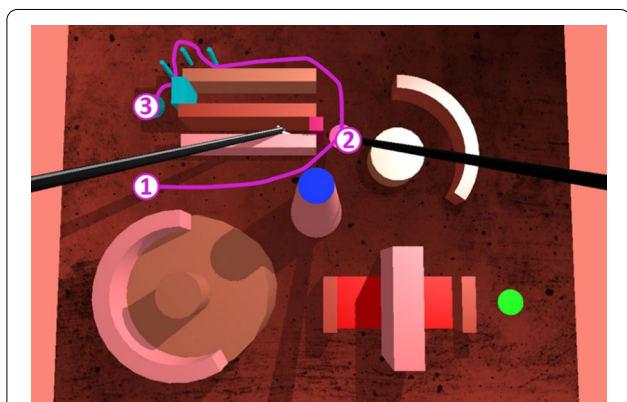


Fig. 3 Screenshot of the trajectory to perform in exercise 2 from the starting position (1) to the ending position (3). The participant, using his/her non-dominant hand had to perform a small trajectory (in purple): **a** start from point 1, **b** go to point 2 that is the first virtual button (located under the label 2) he/she had to push to open the pink virtual gate (on the left of label 2), **c** with his/her dominant hand, pass the latter gate, **d** with the non-dominant hand, slalom from point 2 to point 3 between the 3 light blue poles without touching them, **e** once arrived at point 3, the non-dominant hand had to rotate around its axis (using the instrument wheel) to open the light blue gate and enable the dominant hand to go further between the two upper bars (not represented)

think this simulator could be useful for MIS training?), and finally, three questions were related to the simulator's utility (How do you think this simulator improves the use

of the non-dominant hand?). Finally, the internal consistency test ensured the reliability of the responses [44, 45].

Discriminative validation (relation with experience) and temporal stability

Another approach to validate this simulator was to compare expert and novice performance [46, 47]. The simulator should show that experts perform better than novices. This should indicate that successful performance on the simulator requires surgical expertise. We thus tested here the simulators' discriminative power by comparing participants' performance on the two exercises through three variables (time, number of collisions, and affine velocity). Finally, we tested the temporal stability i.e. the reliability of measurements by a test–retest session. Five participants among the experts' group were randomly selected for this comparison. As recommended, the two testing sessions were separated by at least 1 month without any practice on the simulator [48].

Participants

Seventy seven participants (mean age = 30.8; female = 26; male = 51) agreed to participate in the study. There were 17 experts (mean age = 41.5; female = 4; male = 13) selected regarding the number of medical procedures they already performed (more than 100 interventions as the main operator). Twelve surgeons of intermediate surgery experience (1 female and 11 males, mean age = 28) were selected in the second group as they had between 5 and 20 interventions as the main operator. Twenty-eight inexperienced interns with basic open surgery experience e.g. knot tying and suturing were included in the third group. They additionally observed laparoscopic interventions without performing any operation. There were 8 females and 20 males in this group (mean age = 25.3). Twenty novices unrelated to medical education or surgical skills were also selected, representing 13 females and 7 males (mean age = 31.5). The latter group did not complete the questionnaire because their opinion was not relevant to the content validation process. Among the expert group, one participant did not complete the questionnaire. This experiment did not require any IRB approval. Verbal consent was provided by every participant. Participants' demographic and surgical specialty information is summarized in Table 1.

Data analysis

Data were analyzed using SPSS 21.0 (SPSS Inc.)⁶. To assess content evidence, we applied a one-factor ANOVA to compare groups with post hoc Tukey correction tests.

⁶ See <http://www.ibm.com/software/analytics/spss/products/statistics/index.html>

Table 1 Participants' demographic and specialty

| Group | General information | | Specialty | | | | | | | Experience |
|-----------------------|-------------------------------|------------------|------------|-------------|---------|-----------|----------|-------------|---------|----------------------------------|
| | Total number (Male:Female) | Mean age (range) | Paediatric | Gynaecology | Urology | Digestive | Visceral | Orthopaedic | General | |
| Experts | 17 (13:4) | 41.5 (62–30) | 1 | 1 | 8 | 4 | 2 | 1 | 0 | ≥ 100 interventions ^a |
| Intermediates | 12 (11 :1) | 28 (32–25) | 0 | 0 | 10 | 0 | 2 | 0 | 0 | 5–20 interventions ^a |
| Inexperienced interns | 28 (20 :8) | 25.3 (40–23) | 0 | 0 | 0 | 0 | 0 | 0 | 28 | Weak experience ^b |
| Novices | 20 (7:13) | 31.5 (53–18) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | No experience ^c |

^a As main operator

^b Never performed any laparoscopic operation but had some basic open surgery experience like knot tying and suturing

^c Students with no medical education or surgical skills

A level of $p \leq 0.05$ was considered statistically different. Cronbach alpha test measured the internal consistency of the questionnaire with the α -value set at 0.7). We used non-parametric tests to compare groups regarding time, number of collisions, and affine velocity (discriminative protocol), as the number of samples was too low for parametric tests. The Kruskal–Wallis test compared the three groups, associated with the Kolmogorov–Smirnov for two-by-two comparisons. The statistical threshold was set at $p \leq 0.05$. Finally, correlation coefficients evaluated the temporal stability of the simulator's measure.

Results

The simulator

We designed and prototyped a training simulator using, as possible, affordable components available on the shelf and standard open-source software technologies. Its cost has been evaluated at around \$2500 for a single specimen, including hardware purchases, rapid fabrication, and software development. This cost is compliant with the provision by universities of series of a simulator for free use by students in comparison with the cost of a single high-fidelity simulator (>\$50,000).

Content evidence and its internal consistency

All participants positively rated the simulator whatever the group of inclusion, with mean (\pm SD) scores being 7.25 ± 0.8 , 7.74 ± 0.6 , and 7.72 ± 0.7 for the expert, intermediate and inexperienced levels, respectively. ANOVA did not provide any difference among the three groups ($p = 0.13$, NS). The internal consistency of the content evidence questionnaire revealed a high Cronbach alpha coefficient (0.87).

Discriminative validation (relation with experience) and temporal stability

Exercise 1

Time Experts and intermediates completed this exercise faster than inexperienced interns and novices. Mean times were 82.4 s (SD = 22.1 s), 81.5 s (SD = 23.7 s), 122.3 s (SD = 50 s), and 156.8 s (SD = 53 s) for experts, intermediates, interns, and novices respectively. The Kruskal–Wallis test revealed a significant difference among the groups (Chi-square = 32.80, $p < 0.001$) with a mean rank of 22.41 in the experts, 20.92 in the intermediates, 43.14 in the inexperienced interns, and 58.15 in the novices. Post hoc tests showed significant differences between experts and both inexperienced interns ($p = 0.007$) and novices ($p < 0.001$). Intermediates outperformed inexperienced interns ($p = 0.01$) as well as novices ($p < 0.001$). Finally, inexperienced interns tend to be faster than the novices ($p = 0.06$).

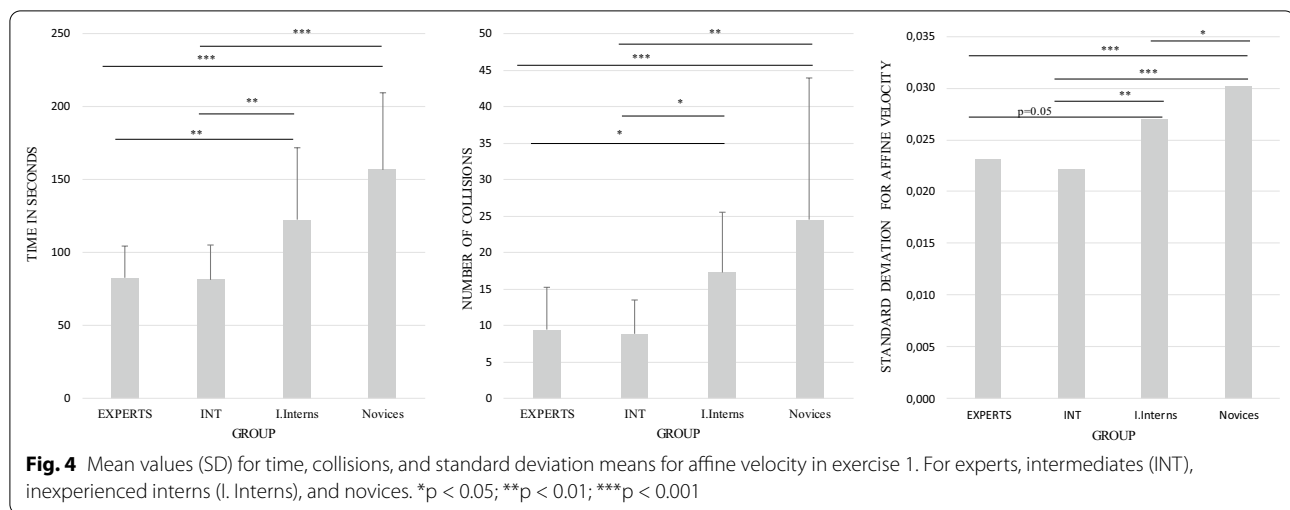
Collisions Experts and intermediates made fewer collisions than inexperienced interns and novices. The mean number of collisions was 9.4 (SD = 5.8), 8.8 (SD = 4.7), 17.3 (SD = 8.2), and 24.5 (SD = 19.5) for experts, intermediates, interns, and novices, respectively. The Kruskal–Wallis test revealed significant differences among groups, (Chi-square = 22.48, $p < 0.001$) with a mean rank of 24.24 in the experts, 22.83 in the intermediates, 46.30 in the inexperienced interns and 51.03 in the novices. Post hoc tests showed significant differences between experts and inexperienced interns ($p = 0.004$), experts and novices ($p < 0.001$) and as well as between intermediates and inexperienced interns ($p = 0.007$); intermediates and novices ($p = 0.001$).

Affine velocity Experts and intermediates outperformed inexperienced interns and novices in the affine velocity variable. Mean affine velocities were 0.023 m/s (SD = 0.0024 m/s), 0.022 m/s (SD = 0.0030 m/s), 0.026 m/s (SD = 0.0055 m/s) and 0.030 m/s (SD = 0.0038 m/s) for experts, intermediates, interns and novices respectively. Kruskal–Wallis test revealed significant differences among groups, (Chi-square = 30.65, $p < 0.001$) with a mean rank of 25.76 for the experts, 19.08 for the intermediates, 41.86 for the inexperienced interns, and 58.20 for the novices. Post hoc tests showed significant differences between experts and novices ($p < 0.001$) and between experts and inexperienced interns ($p = 0.05$). Post hoc tests showed also significant differences between intermediates and inexperienced interns ($p = 0.009$) as well as between intermediates and novices ($p < 0.001$). There was also a significant difference between inexperienced interns and novices ($p = 0.003$).

Exercise 2

Time Experts and intermediates completed the second exercise faster than inexperienced interns and novices. The mean movement time was 109.5 s (SD = 30 s), 107.1 s (SD = 34.1 s), 147.3 s (SD = 41 s), and 188.1 s (SD = 46.6 s) for experts, intermediates, interns, and novices respectively. The Kruskal–Wallis test revealed a significant difference among the groups, (Chi-square = 34.09, $p < 0.001$) with a mean rank of 21.97 in the experts, 21.88 in the intermediates, 42.09 in the inexperienced interns, and 59.43 in the novices. Post hoc tests showed significant differences between experts and inexperienced interns ($p = 0.01$); experts and novices ($p < 0.001$) but also between intermediates and inexperienced interns ($p = 0.02$) as well as between intermediates and novices ($p < 0.001$). Movement time was also different between inexperienced interns and novices ($p = 0.02$).

Collisions Experts and intermediates made fewer collisions than inexperienced interns and novices. The mean number of collisions was 17.7 (SD = 10.3), 19.6



(SD = 5.9), 23.3 (SD = 13), and 37.5 (SD = 20.1) in experts, intermediates, interns, and novices respectively. The Kruskal–Wallis test revealed significant differences among groups, (Chi-square = 14.703, $p = 0.002$) with a mean rank of 27.21 in the experts, 33.88 in the intermediates, 37.52 in the inexperienced interns, and 54.18 in the novices. Post hoc tests showed significant differences between experts and novices ($p = 0.002$) and between intermediates and novices ($p = 0.03$). There was also a significant difference between inexperienced interns and novices ($p = 0.03$)

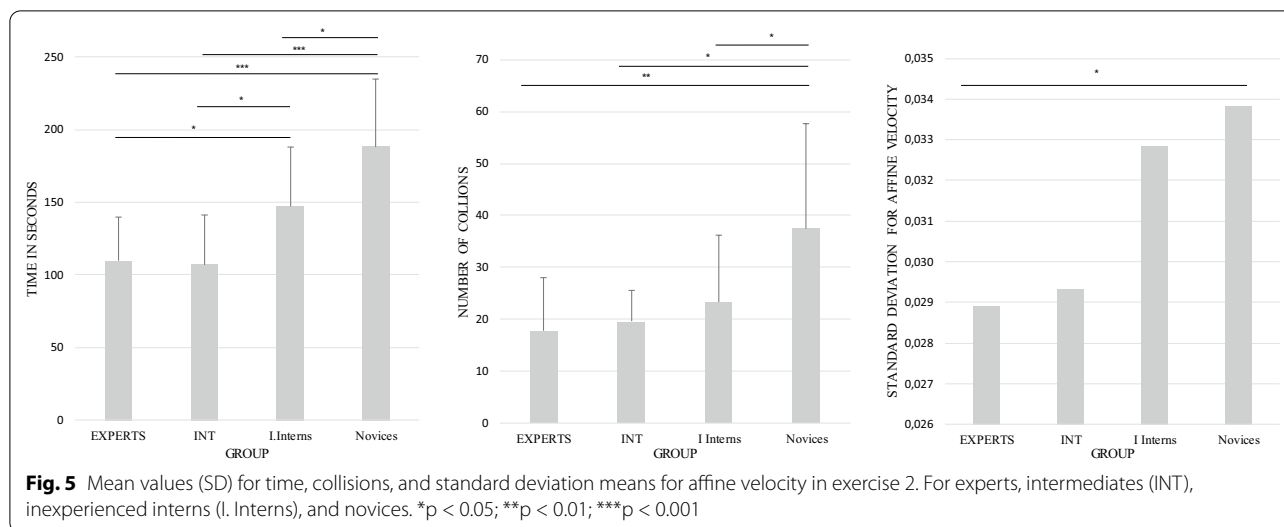
Affine velocity Experts and intermediates outperformed inexperienced interns and novices. Mean affine velocity were 0.028 m/s (SD = 0.0042 m/s), 0.029 m/s (SD = 0.0055 m/s), 0.032 m/s (SD = 0.0056 m/s) and 0.033 m/s (SD = 0.0051 m/s) in experts, intermediates, interns, and novices, respectively. The Kruskal–Wallis test revealed significant differences among groups, (Chi-squared = 10.41, $p = 0.015$) with a mean rank of 26.88 in the experts, 31.83 in the intermediates, 42.89 in the inexperienced interns and 48.15 in the novices. Post hoc tests showed significant differences between experts and novices ($p = 0.04$) and tend to be different between experts and inexperienced interns ($p = 0.06$).

Finally, regarding the temporal stability of the simulator's measure, we found a correlation for the number of collisions and for the time variables between the two sessions ($r = 0.89$; $p = 0.04$).

Discussion

This study demonstrated the reliability and content evidence of this low-fidelity simulator in laparoscopic motor skills learning. The difference among groups' performance accounts for the discriminative power of the

simulator. The users uniformly rated the way the simulator exhibits actual motor abilities required by real surgery. Surgeons' global agreement about the usefulness of the simulator for surgical curricula is attested through a large number of participants, where three levels of practice were represented (from inexperienced interns to expert surgeons, with an intermediate level). The outcome scores are in line with the outcomes from other studies dealing with the validation of computer-based systems [49]. For instance, the MIST VR (MenticeMedical Simulation, Gothenburg, Sweden) was rated by an average score of 7 in a comparable study [49]. Since this validation process, several studies used the MIST VR to train basic laparoscopic skills, thereby demonstrating the relevance of the simulator in surgery training [47, 50]. Taken together, the surgeons validated positively how laparoscopic skills are brought into play and assessed in the simulator. Regarding its discriminative power, both exercises exhibited a slightly different pattern of results. In general, we observed that experts and intermediates outperformed inexperienced interns and novices according to the three dependent variables we selected in both exercises, i.e. movement time duration, number of collisions, and affine velocity (Figs. 4 and 5). This confirmed that the simulator discriminated among four expertise levels although to a lesser extent between inexperienced interns and novices. The simulator thus highlighted the main abilities needed by MIS. When looking closer to the results of exercise 1, experts and intermediates generally performed at the same level. This outcome is consistent as both groups had real experience and this first exercise did not require advanced laparoscopic skills. The same argument can explain the discrepancy between these groups and the two others, which did not master laparoscopic motor skills, even basic ones. This is a positive



outcome as our purpose was to engineer a low-fidelity simulator for the training of basic surgical skills, thereby more fitted for beginners [13]. In the second exercise, the pattern of results was quite similar to that of the first exercise. However, when looking closer at data (Fig. 5), only the novice group is systematically outperformed by the other groups (except for affine velocity where only the experts outperformed the novices). To better analyze and interpret these results, we should consider that the task was bi-manual. Surgery skills required bi-manual coordination [51] and this ability is essential in laparoscopic surgery [7, 52] and open surgery [5]. Thus, prior open surgical experience helped the experts, intermediates, and inexperienced interns to overcome exercise 2, while the novices did not benefit from previous experience. As a consequence, a laparoscopic simulator should include training for bi-manual tasks.

Motor skills learning is a rapid process when training begins. Then, it slows down to tend towards an asymptote, in reference to a classic learning curve. This is probably why it was difficult to clearly distinguish performance between expert and average surgeons [53]. They were categorized on the basis of professional experience duration and this is not a discriminative factor when testing basic skills. In relation to sporting skills [54], we can keep this hypothesis to be further tested. Expertise is a long process integrating speed, accuracy, and economy of resources, together constituting efficiency. As early as 1995, DesCôteaux and Leclère [55] considered that “*visuospatial perceptual skills (the ability to represent mentally the physical environment and the movement to be performed) are the major determinants of surgical technical performance*”. From this view point, our hypothesis would be tested by integrating other variables associated

with a surgical experience, specifically by making the simulator testing specific skills, regarding surgical specialties.

Testing the simulator through one single trial may be seen as a limitation. However, running several trials would include the risk of habituation, i.e. starting a fast learning stage, as observed by Dayan and Cohen [56]. This can potentially affect the outcomes, particularly in the novices and inexperienced interns. Despite only one trial was performed, thus guaranteeing that no learning process could occur, performance between inexperienced interns and novices was often comparable and the difference was weak when reaching significance (see variable “time” during exercise 1 in “Results” section). This shows that medical experience with the mastery of specific skills is not generalized to surgery. Therefore, this should prompt to form only one group of novices in future experiences [56].

Taken together, the results nevertheless suggest that the three variables were relevant. These should be analyzed in a complementary way to provide an objective overview of the trainee’s skills. This confirms the existing literature, stating that time and collision are indirect indicators of laparoscopic proficiency, needed to be associated with motion characteristics, provided by the affine velocity [15, 36].

The evolution of the simulator could incorporate new skills such as suturing and knot tying with the aim to test more advanced techniques. As previously proposed, introducing specific surgical techniques remains an open option. This would require additional software development, however without hardware changes. This would also require coding new exercises in C++, based on the Chai3D library. From our own experience, we estimate that the conception of a single new exercise testing

a new skill would require around one person.month of work.

Conclusion

The outcomes showed that integrating the affine velocity metrics is a reliable tool according to both inquiry results and statistical tests. Also, the simulator exhibits innovative devices providing easy testing through several games and autonomic assessing of the main laparoscopic basic skills. It provides objective experience discrimination which is lacking in low-fidelity simulators [14, 15, 19], for a low additional cost compared to high-fidelity simulators [22]. It thus proposes a pragmatic trade-off between functionalities and cost. However, this probably needs to be confirmed by investigating the implementation of this simulator in an early training stage along with considering the impact of haptic feedback on learning, and the validation of the discriminative power of affine velocity on more difficult exercises.

Abbreviations

ANOVA: Analysis of variance; IRB: Institutional Review Board; MIS: Minimally Invasive Surgery; NS: Non significant; SD: Standard deviation.

Acknowledgements

The authors would like to thank the ANR for financing SAMSEI project.

Authors' contributions

BDW and CB actively designed and built the prototype of this simulator. RM, AL, and NH supervised this step and conducted the experimental sessions with BDW and CB. XM and CC oversaw the experimental design and the writing/revising of this article. All authors contributed equally to the writing and revising of this paper. All authors read and approved the final manuscript.

Funding

Funding has been provided by the ANR (French National Agency for Research) in the context of the "Learning Strategies for Healthcare Professions in Immersive Environment" (SAMSEI) project (ANR-11-IDFI-0034) under the supervision of Pr. X. Martin, in support and conjunction with INSA Lyon and Lyon 1 University. This funding covered the prototyping costs, a Ph.D. salary, and the current scientific publication. This study report has been provided to SAMSEI which let the authors freely build this experimental protocol and has accepted that experimental data be published.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on a reasonable request.

Declarations

Ethics approval and consent to participate

This experimental design is a behavioral study based on observation. It is a non-interventional experiment as defined and described by the French law Jardé. This means that neither intrusive intervention in the participants' organism, nor ingestion of active molecules present in a drug have been carried out. In line with the Helsinki declaration, all participants were instructed about the main aims of the experiment, and the main operations they will have to manage. After that, all participants signed a written consent describing their rights. In compliance with French ethical law and with the Helsinki declaration, we did not submit a file to an ethics committee. All informed consents are available on request.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Inter-University Laboratory of Human Movement Science (EA 7424), Univ Lyon, University Claude Bernard Lyon 1, 27-29 Boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex, France. ² INSA Lyon, Ampère (UMR5005), Univ Lyon, 25 av. Jean Capelle ouest, 69621 Villeurbanne Cedex, France. ³ Faculty of Medicine, Surgery School, Univ Lyon, University Claude Bernard Lyon 1, 8 Avenue Rockefeller, 69003 Lyon, France. ⁴ Service d'Urologie et de chirurgie de la Transplantation, Hôpital Édouard Herriot, Hospices Civils de Lyon, 5 Place d'Arsonval, 69003 Lyon, France.

Received: 14 September 2020 Accepted: 1 March 2021

Published online: 18 March 2021

References

- Roumm AR, Pizzi L, Goldfarb NI, Cohn H. Minimally invasive: minimally reimbursed? an examination of six laparoscopic surgical procedures. *Surg Innov*. 2005;12:261–87. <https://doi.org/10.1177/155335060501200313>.
- Sheetz KH, Norton EC, Birkmeyer JD, Dimick JB. Provider experience and the comparative safety of laparoscopic and open colectomy. *Health Serv Res*. 2017;52(1):56–73. <https://doi.org/10.1111/1475-6773.12482>.
- Spanjersberg WR, van Sambeek JDP, Bremers A, Rosman C, van Laarhoven CJHM. Systematic review and meta-analysis for laparoscopic versus open colon surgery with or without an eras programme. *Surg Endosc*. 2015;29:3443–53. <https://doi.org/10.1007/s00464-015-4148-3>.
- Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and technical skills performance in health care: a systematic review. *Med Educ*. 2015;49(11):1065–85. <https://doi.org/10.1111/medu.12786>.
- Molinas CR, Binda MM, Campo R. Dominant hand, non-dominant hand, or both? the effect of pre-training in hand-eye coordination upon the learning curve of laparoscopic intra-corporeal knot tying. *Gynecol Surg*. 2017;14(1):12. <https://doi.org/10.1186/s10397-017-1015-3>.
- Yoshida S, Fukuyo T, Saito K, Kihara K, Fujii Y. Real-time three-dimensional image angle rectification to improve hand-eye coordination in single-port laparoendoscopic surgery. *Int J Urol*. 2017;24(8):639–40. <https://doi.org/10.1111/iju.13371>.
- Nepomnayshy D, Alseidi AA, Fitzgibbons SC, Stefanidis D. Identifying the need for and content of an advanced laparoscopic skills curriculum: results of a national survey. *Am J Surg*. 2016;211(2):421–5. <https://doi.org/10.1016/j.amjsurg.2015.10.009>.
- Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *N Engl J Med*. 2006;355(25):2664–9. <https://doi.org/10.1056/nejmra054785>.
- Badash I, Burt K, Solorzano CA, Carey JN. Innovations in surgery simulation: a review of past, current and future techniques. *Ann Transl Med*. 2016;4(23):453–453. <https://doi.org/10.21037/atm.2016.12.24>.
- Amitai Ziv PW, Stephen D. Small: patient safety and simulation-based medical education. *Med Teach*. 2000;22(5):489–95. <https://doi.org/10.1080/01421590050110777>.
- Dunkin B, Adrales GL, Apelgren K, Mellinger JD. Surgical simulation: a current review. *Surg Endosc*. 2006;21(3):357–66. <https://doi.org/10.1007/s00464-006-9072-0>.
- Schmitt F, Mariani A, Eyssartier E, Granry J-C, Podevin G. Learning laparoscopic skills: observation or practice? *J Laparoendosc Adv Surg Tech*. 2018;28(1):89–94. <https://doi.org/10.1089/lap.2017.0254>.
- Munshi F, Lababidi H, Alyousef S. Low- versus high-fidelity simulations in teaching and assessing clinical skills. *J Taibah Univ Med Sci*. 2015;10(1):12–5. <https://doi.org/10.1016/j.jtumed.2015.01.008>.
- Laguna MP, de Reijke TM, Wijkstra H, de la Rosette J. Training in laparoscopic urology. *Curr Opin Urol*. 2006;16(2):65–70. <https://doi.org/10.1097/01.mou.0000193377.14694.7f>.
- Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM. Virtual reality simulation for the operating room. *Ann Surg*. 2005;241(2):364–72. <https://doi.org/10.1097/01.sla.0000151982.85062.80>.

16. Mulla M, Sharma D, Moghul M, Kailani O, Dockery J, Ayis S, Grange P. Learning basic laparoscopic skills: a randomized controlled study comparing box trainer, virtual reality simulator, and mental training. *J Surg Educ*. 2012;69(2):190–5. <https://doi.org/10.1016/j.jvsurg.2011.07.011>.
17. Maran NJ, Glavin RJ. Low- to high-fidelity simulation—a continuum of medical education? *Med Educ*. 2003;37(s1):22–8. <https://doi.org/10.1046/j.1365-2923.37.s1.9.x>.
18. Parmar S, Delaney CP. The role of proximate feedback in skills training. *Surgeon*. 2011;9:26–7. <https://doi.org/10.1016/j.surge.2010.11.006>.
19. Hammoud MM, Nuthalapaty FS, Goepfert AR, Casey PM, Emmons S, Espey EL, Kaczmarczyk JM, Katz NT, Neutens JJ, Peskin EG. To the point: medical education review of the role of simulators in surgical training. *Am J Obstet Gynecol*. 2008;199(4):338–43. <https://doi.org/10.1016/j.ajog.2008.05.002>.
20. McKenna DT, Mattar SG. What is wrong with the training of general surgery? *Adv Surg*. 2014;48(1):201–10. <https://doi.org/10.1016/j.yasu.2014.05.010>.
21. Crochet P, Aggarwal R, Knight S, Berdah S, Boubli L, Agostini A. Development of an evidence-based training program for laparoscopic hysterectomy on a virtual reality simulator. *Surg Endosc*. 2016;31(6):2474–82. <https://doi.org/10.1007/s00464-016-5249-3>.
22. Li MM, George J. A systematic review of low-cost laparoscopic simulators. *Surg Endosc*. 2016;31(1):38–48. <https://doi.org/10.1007/s00464-016-4953-3>.
23. Stefanidis D, Acker C, Heniford BT. Proficiency-based laparoscopic simulator training leads to improved operating room skill that is resistant to decay. *Surg Innov*. 2008;15(1):69–73. <https://doi.org/10.1177/1553350608316683>.
24. Yiasemidou M, de Siqueira J, Tomlinson J, Glassman D, Stock S, Gough M. “take-home” box trainers are an effective alternative to virtual reality simulators. *J Surg Res*. 2017;213:69–74. <https://doi.org/10.1016/j.jss.2017.02.038>.
25. Leijte E, Arts E, Witteman B, Jakimowicz J, Blaauw ID, Botden S. Construct, content and face validity of the eoSIm laparoscopic simulator on advanced suturing tasks. *Surg Endosc*. 2019;33(11):3635–43. <https://doi.org/10.1007/s00464-018-06652-3>.
26. Poursartip B, LeBel M-E, Patel RV, Naish MD, Trejos AL. Analysis of energy-based metrics for laparoscopic skills assessment. *IEEE Trans Biomed Eng*. 2018;65(7):1532–42. <https://doi.org/10.1109/tbme.2017.2706499>.
27. Schout BMA, Hendriks AJM, Scheele F, Bemelmans BLH, Scherpbier AJJA. Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surg Endosc*. 2009;24(3):536–46. <https://doi.org/10.1007/s00464-009-0634-9>.
28. AERA, APA, NCME. The standards for educational and psychological testing. Technical report, American Psychological Association. 2014. <https://www.apa.org/science/programs/testing/standards>.
29. Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med*. 2006;119(2):166–716616. <https://doi.org/10.1016/j.amjmed.2005.10.036>.
30. Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *J Gastrointest Surg*. 2007;12(2):210–2. <https://doi.org/10.1007/s11605-007-0355-0>.
31. Hedman L, Ström P, Andersson P, Kjellin A, Wredmark T, Felländer-Tsai L. High-level visual-spatial ability for novices correlates with performance in a visual-spatial complex surgical simulator task. *Surg Endosc*. 2006;20(8):1275–80. <https://doi.org/10.1007/s00464-005-0036-6>.
32. Young JQ, Merrienboer J.V, Durning S, Cate O.T. Cognitive load theory: implications for medical education: AMEE guide no. 86. *Med Teach*. 2014;36(5):371–84. <https://doi.org/10.3109/0142159x.2014.889290>.
33. Mayer RE, Moreno R. Nine ways to reduce cognitive load in multimedia learning. *Educ Psychol*. 2003;38(1):43–52. https://doi.org/10.1207/s15326985ep3801_6.
34. Tendick F, Downes M, Goktekin T, Cavusoglu M.C., Feygin D, Wu X, Eyal R, Hegarty M, Way L.W. A virtual environment testbed for training laparoscopic surgical skills. *Presence Teleoperators Virtual Environ*. 2000;9(3):236–55. <https://doi.org/10.1162/105474600566772>.
35. Kozak I, Banerjee P, Luo J, Luciano C. Virtual reality simulator for vitreoretinal surgery using integrated OCT data. *Clin Ophthalmol*. 2014. <https://doi.org/10.2147/ophth.s58614>.
36. Spruit EN, Band GPH, Hamming JF, Ridderinkhof KR. Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. *Psychol Res*. 2013;78(6):878–91. <https://doi.org/10.1007/s00426-013-0525-5>.
37. Blackhall VI, Cleland J, Wilson P, Moug SJ, Walker KG. Barriers and facilitators to deliberate practice using take-home laparoscopic simulators. *Surg Endosc*. 2018;33(9):2951–9. <https://doi.org/10.1007/s00464-018-6599-9>.
38. Lahanas V, Loukas C, Georgiou K, Lababidi H, Al-Jaroudi D. Virtual reality-based assessment of basic laparoscopic skills using the leap motion controller. *Surg Endosc*. 2017;31(12):5012–23. <https://doi.org/10.1007/s00464-017-5503-3>.
39. Morasso P. Spatial control of arm movements. *Exp Brain Res*. 1981;42(2):223–7. <https://doi.org/10.1007/BF00236911>.
40. Pollick FE, Sapiro G. Constant affine velocity predicts the 1/3 power law of planar motion perception and generation. *Vis Res*. 1997;37(3):347–53. [https://doi.org/10.1016/S0042-6989\(96\)00116-2](https://doi.org/10.1016/S0042-6989(96)00116-2).
41. Maoz U, Berthoz A, Flash T. Complex unconstrained three-dimensional hand movement and constant equi-affine speed. *J Neurophysiol*. 2009;101(2):1002–15. <https://doi.org/10.1152/jn.90702.2008>.
42. Cifuentes J, Boulanger P, Pham MT, Moreau R, Prieto F. Automatic gesture analysis using constant affine velocity. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society. IEEE, Chicago, Illinois, USA. 2014. <https://doi.org/10.1109/embc.2014.6943964>.
43. Carter FJ, Schijven MP, Aggarwal R, Grantcharov T, Francis NK, Hanna GB, Jakimowicz JJ. Consensus guidelines for validation of virtual reality surgical simulators. *Surg Endosc*. 2005;19(12):1523–32. <https://doi.org/10.1007/s00464-005-0384-2>.
44. Anastasi A. Psychological testing. 4th ed. Oxford: Macmillan; 1976. p. 750750.
45. Cronbach LJ. Essentials of psychological testing. *Sci Educ*. 1951;35(4):227–227. <https://doi.org/10.1002/sce.3730350432>.
46. Mathis KL, Wiegmann DA. Construct validation of a laparoscopic surgical simulator. *Simul Healthc J Soc Simul Healthc*. 2007;2(3):178–82. <https://doi.org/10.1097/sih.0b013e318137aba1>.
47. Gallagher AG, Lederman AB, McGlade K, Satava RM, Smith CD. Discriminative validity of the minimally invasive surgical trainer in virtual reality (MIST-VR) using criteria levels based on expert performance. *Surg Endosc*. 2004;18(4):660–5. <https://doi.org/10.1007/s00464-003-8176-z>.
48. Vallerand RJ. Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: implications pour la recherche en langue française. *Can Psychol*. 1989;30(4):662–80. <https://doi.org/10.1037/h0079856>.
49. Maithe S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D. Construct and face validity of MIST-VR, endotower, and CELTS. *Surg Endosc*. 2005;20(1):104–12. <https://doi.org/10.1007/s00464-005-0054-4>.
50. Aggarwal R, Grantcharov T, Moorthy K, Hance J, Darzi A. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg*. 2006;191(1):128–33. <https://doi.org/10.1016/j.amjsurg.2005.10.014>.
51. Tchantchaleishvili V, Myers PO. Hand laterality and acquired ambidexterity in surgical training. *Ann Surg*. 2016;264(6):18–9. <https://doi.org/10.1097/sla.0000000000001951>.
52. Anna S, Gregory A, Margaret M, Marc S. Ambidexterity in laparoscopic surgical skills training. *Stud Health Technol Inform*. 184(Medicine Meets Virtual Reality 20). 2013;412–416. <https://doi.org/10.3233/978-1-61499-209-7-412>.
53. Fritz T, Stachel N, Braun BJ. Evidence in surgical training—a review. *Innov Surg Sci*. 2019;4:7–13. <https://doi.org/10.1515/iss-2018-0026>.
54. Cocks M, Moulton C-A, Luu S, Cil T. What surgeons can learn from athletes: mental practice in sports and surgery. *J Surg Educ*. 2014;71:262–9. <https://doi.org/10.1016/j.jvsurg.2013.07.002>.
55. DesCôteaux JG, Leclère H. Learning surgical technical skills. *Can J Surg*. 1995;38:33–8.
56. Dayan E, Cohen L. Neuroplasticity subserving motor skill learning. *Neuron*. 2011;72(3):443–54. <https://doi.org/10.1016/j.neuron.2011.10.008>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.