RESEARCH

A Haptic laparoscopic trainer based on Affine Velocity Analysis: Engineering and preliminary Results

Benjamin De Witte[1†], Charles Barnouin[2†], Richard Moreau[2], Arnaud Lelevé[2*], Xavier Martin[3,4], Christian Collet[1] and Nady Hoyek[1]

*Correspondence: arnaud.leleve@insa-lyon.fr
[2]Univ Lyon, INSA Lyon, Ampère (UMR5005), 25 av. Jean Capelle ouest, 69621 Villeurbanne Cedex, FR
Full list of author information is available at the end of the article
†Equal contributor

Abstract

Background: General agreement exists upon the importance of acquiring laparoscopic skills outside the operation room through simulation-based training. Nevertheless such a learning is hindered by several flaws. High-fidelity simulators are cost-prohibitive (which limits training opportunities), and elicit a high cognitive load. Low-fidelity simulators lack in haptic, direct, and summative feedback. This paper introduces a new low-fidelity simulator integrating effective learning features and a new assessment variable while limiting the associated costs. We also detail primary validation results.

Methods: We engineered a low fidelity laparoscopic simulator taking into account psycho-motor skills, direct and summative feedback, and engineering key features (haptic feedback and complementary assessment variables).
77 participants with 4 different surgical skill levels (17 experts; 12 intermediates; 28 inexperienced interns and 20 novices) tested the simulator. We checked the content validity using a 10-point Likert scale and the discriminative power by comparing the 4 groups' performance over two sessions, using 3 variables: time, number of collisions, and affine velocity.

Results: The mean value score of the content validation was 7.57/10. The statistical analysis yielded performance discrepancies on the selected variables among the groups ($p<0.001$).

Conclusion: We developed an affordable and validated simulator for testing and learning basic laparoscopic skills. The results exhibit three levels of performance. Results show that the embedded evaluation variables are complementary and provide realistic results. The inclusion of a new variable and, meanwhile, haptic, direct, and summative feedback is innovative regarding low-fidelity simulators. Limitations and implementation conditions of the simulator in the surgical curricula are discussed.

Keywords: Minimal Invasive Surgery; Motor Skills; Assessment metrics; Affine Velocity; Simulator validation

Introduction

Minimally Invasive Surgery (MIS) has brought more comfort for the patients in comparison to open surgery [1, 2, 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, a different skill set, e.g. involving spatial abilities [4], bi-manual coordination [5], and hand-eye coordination [6], is essential in MIS practice. At the early stages of MIS learning, the relevance of these skills is increased [7]. How these skills are learned is one of the major concerns of young surgeons training [8]. Although simple observation of laparoscopic skills may help to learn, direct practice on the virtual reality trainer ensures more effective training [9]. During the past two decades, simulation-based training and simulators have been more extensively used in surgeons' training [10], without the risk of harming patients [11, 12], and with health care cost reductions [11]. Two kinds of simulators coexist: low and high fidelity simulators. Low fidelity ones recreate laparoscopic working conditions with simple instruments. They are better suited to novice learners [12] as they are efficient for basic skills acquisition [13] such as clipping, grasping, or cutting. Indeed, a simple simulation training without the pressure of result and enabling to progress at self-pace is less likely to overwhelm novice learners [14]. However, low-fidelity simulators have obvious limitations. First, they do not recreate the human body, making the illusion of reality impossible [15]. Second, they fail to provide immediate feedback when an error occurs, or summative feedback on completion of a trial [15, 16], path length or economy of movements either [15, 16, 17]. Inadequate or insufficient feedback seriously hinders learning as their provision is fundamental for effective learning [12, 18, 19]. Finally, the use of low fidelity simulators requires the direct observation of an expert for full assessment [15, 20]. Usually few available, these experts do not operate patients while they monitor such hands-on training sessions, and do extra hours that are costly for the Society. Furthermore, their assessment is inherently subjective. This reduces effective training opportunities for novice surgical students [21]. Once one's proficiency is developed, high fidelity simulators are relatively more effective.

As they can replicate a whole procedure [10] (hysterectomy [22], cholecystectomy [23], or salpingectomy [24], for instance), they are more suited to improve specific procedural skills. As an example, training with the LAP Mentor$^{\text{TM}}$ (Simbionix Corporation, Cleveland, Ohio, USA) or the LAPSIM® (Surgical Science, Gothenburg, Sweden) improves the performance of complete and complex laparoscopic procedures [10, 22, 23]. Nevertheless, high fidelity simulators are very cost-prohibitive and may not be accessible for regular and personal use [20, 25]. Besides, this lack of availability limits ongoing training and hinders thereby skill retention [26]. Finally, as shown in aviation training, the knowledge required to perform the task, added to the complexity of the instrumentation and the pressure to perform may overwhelm novice users [27]. The latter arguments show that high-fidelity simulators are better suited for almost graduated interns than for novices. Globally, training skills on both types of simulators are transferable to actual surgery conditions [28, 29]. Nevertheless, a recent study showed that a 5-day training course using a box-trainer has better outcomes than the same training using a high-fidelity simulator [30]. Even if low-fidelity and high-fidelity simulators may be considered as a continuum in the learning process [13, 18], a gap remains between them: on the one hand, high-fidelity simulators provide accurate summative and haptic feedback but are expensive, overloaded and barely available. On the other hand, low-fidelity simulators cost less and are broadly used by surgical students [17], but require supervision and do not provide enough feedback for effective learning at a stage where it is essential to acquire the appropriate skills. To illustrate this gap, Yiasemidou et. al. [29] showed that the students who autonomously trained on a “take-home” box trainer (≈\$500) at home during 6 weeks performed better than the ones who practiced on the High-Fidelity Virtual-Reality-based simulator (≈\$70,000). Therefore, a solution mixing low-cost, broad availability, autonomous use, and objective automatic assessment should efficiently fill this gap. Concerning the 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 [14, 16]. A review of existing metrics is provided in [31], which 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 [32]. To illustrate this situation, experiments described in [33] revealed that experts performed more smooth, accurate and fine movements compared to novices. Usual metrics do not directly evaluate the curvature of instrument motions, which helps reveal the level of expertise of users, as experimentally shown in this study. To address the above-mentioned flaws, this study aims to develop and validate a new low-cost (≈$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 [34]), ii) both haptic and summative feedback, and iii) a new assessment variable (in a similar approach as Poursartip et al. [31]) permitting the evaluation of the delicateness of motions. Thus, this training simulator could be intensively used by students in full autonomy. As shown by Shout et al. [35], before testing the training efficiency of a simulator, scientific validation needs to be processed. According to current standards for educational and psychological testing [36, 37], we present the preliminary validation results (reliability, content, and relation with experimental designs).

Material and 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 [38]. 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, 5 obstacles are displayed along the trajectory, each one requiring the execution of a specific action (Figure 1). The exercises were submitted to and validated by the Head of Lyon Surgery Department. They combine memory work, skills involving visual-spatial ability [39]. Considering the high cognitive load involved by high-fidelity simulators, which may negatively impact the ability to learn medical gestures [40], we followed the recommendations of Mayer and Moreno [41] and kept the exercises and the virtual environment simple, focusing on the gestures (trajectories and tool synchronization).

Insert figure 1 around here

To avoid high computing costs, we used open-source software and affordable material [42]. 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 [43]. We developed the software with Microsoft Visual Studio® 2015[1] and we used the open-source Haptic Framework CHAI3D 3.1.1[2] to make the haptic devices, the 3D virtual world, and the Open-GL 4.4[3] renderer communicating to each other. We designed all vir-

[1]see `https://www.microsoft.com/france/visual-studio/`
[2]see `http://www.chai3d.org/`
[3]see `https://www.opengl.org/`

tual objects with the 3DS Max software 2016[4]. We finally collected and processed performance data (affine velocity, motion duration, numbers of collisions) through Matlab® R2015a (Mathworks)[5].

Insert figure 2 around here

Affine velocity is a metric that takes the geometry of the trajectory into account, as well as its dynamics. In 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 [44]. Furthermore, this relationship conforms to an empirical two-thirds power law [45]. For a MIS-tool 3D trajectory, this property had to be adapted for spatial motion. Some experimental results suggested that the two-third power law does not fit 3D motions and that a one-sixth power law is needed [46]. The relation between the Euclidean velocity v, the curvature κ, and the torsion τ is defined as:

$$v = v_a.\kappa.\alpha.|\tau|.\beta \qquad (1)$$

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 [47], 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 (tra-

[4] see `http://www.autodesk.fr/products/3ds-max/overview`
[5] see `https://fr.mathworks.com`

jectories, 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 [47], 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):

$$\log(v) = \log(v_a) + \alpha . \log(\kappa) + \beta . \log(|\tau|) \tag{2}$$

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 gestures to allow a realistic computation of affine velocity.For this medical gesture, we found α = -0.048 and $\beta = -0.0026$.

Validation processes

According to the American Psychological Association guidelines [48] and standards [36], we validated the simulator's content evidence and discriminative power.

Content Evidence validation and its internal consistency

We examined if the simulators' items completely represent a basic surgical skill-learning tool. For the first exercise (figure 1), 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, 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, 5 the didactic components, 4 the perception of reliability, 7 questions were linked to the training and learning power, and finally, 3 questions were related to the simulators utility. Finally, the internal consistency test ensured the reliability of the responses [49, 50].

Discriminative validation (relation with experience) and temporal stability

Another approach to validate this simulator was to compare expert and novice performances [51, 52]. Obviously, 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' performances 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 one month without any practice on the simulator [53].

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 with reference to 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 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.

Insert Table 1 around here

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. 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 is too low for parametric tests. We performed the Kruskal-Wallis test to compare the three groups, and 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

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$\pm$0.8, 7.74$\pm$0.6, and 7.72$\pm$0.7 for the expert,

[6] See `http://www.ibm.com/software/analytics/spss/products/statistics/index.html`

intermediate and inexperienced levels, respectively. ANOVA did not provide any difference among the 3 groups (p=0.13, NS). The internal consistency of the content evidence questionnaire revealed a high Cronbach alpha coefficient of 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 = 32.802, 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 unexperienced 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).

Insert figure 3 around here

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).

Insert figure 4 around here

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 we observed among groups' performance also 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 3 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 [54]. The MIST VR (MenticeMedical Simulation, Gothenburg, Sweden), was for instance rated by an average score of 7 in a comparable study [54]. Since this validation process, several studies used the MIST VR to train basic laparoscopic skills, thereby demonstrating the relevance of the simulator in laparoscopic training [52, 55]. Taken together, this indicates that he surgeons validated positively the way in which laparoscopic skills are simulated. Regarding the discriminative power of the simulator, both exercises exhibited a slightly different pattern of results. In general, we can observe that for both exercises, experts and intermediates outperformed inexperienced interns and novices according to the three dependent variables we selected i.e. movement time duration, number of collisions, and affine velocity (figures 1 and 2). This confirmed that the simulator discriminated among 4 expertise levels. In other words, surgical skills are required to perform well on the simulator. The simulator thus highlighted the main abilities needed by MIS. When looking closer to the results of exercise 1, experts and intermediates perform similarly. This outcome is consistent as these groups have real experience and this first exercise did not require advanced laparoscopic skills. The same argument can explain the discrepancy between experts, intermediates, and the other groups, which do 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 is quite similar to that of the first exercise. However, when looking closer at data (figure 3), only the novice group is systematically outperformed by the other groups (with the exception of affine velocity where only the experts outperformed the novices). To better analyze and understand these results, we should consider that this was a bi-manual task, while the first exercise only involved one hand. Surgery skills required the coordination of the two hands [56] and the ability of using both hands is essential in laparoscopic surgery [7, 57] and in open surgery [5]. Thus, prior open surgical experience helped the experts, intermediates and inexperienced interns to overcome the difficulty of the bimanual task, while the novices did not benefit from previous experience. A laparoscopic simulator should obviously include bi-manual tasks. While analyzing the main outcomes, we can underline some limitations. The discriminative validation protocol may also be seen as a limitation when considering that one single trial was performed. Results would have been reinforced with several trials in each experimental condition. However, running several trials means taking the risk of starting a fast learning stage [58], which can also affect the outcomes. Indeed, Dayan and Cohen observed that a fast learning stage can improve motor execution after one repetition [58]. However, taken together, the results of the validation protocols suggest that the three variables we selected to evaluate the laparoscopic proficiency on this simulator are relevant. Indeed, for an objective overview of the trainee's skills, the three variables should be analyzed in a complementary way. This last point meets the existing literature, stating that time and collisions are indirect indicators of laparoscopic proficiency and that they need to be associated with objective measures i.e. motion characteristics, provided by the affine velocity [14, 16].

Conclusion

The outcomes of the aforementioned validation protocols showed that the simulator we described is a reliable and innovative device reproducing the main laparoscopic basic motor skills. It required bi-manual coordination while providing haptic and summative feedback. The inclusion of haptic feedback and complementary (innovative) evaluation variables such as the affine velocity, represents originality regarding low-fidelity simulators [15, 16, 20]. The low associated costs (≈\$2500) obviously contrast with high-fidelity simulators [25]. Its virtual operation theater (unlike Islam et al. in [59]) enables the development of new sets of perfectly reproducible exercises featuring objective instantaneous haptic and summative feedback, thus contrasting with usual low-fidelity simulators. As training solutions outside the operation room need to be engineered for the initial training period of surgeons [5, 60], and as this simulator main aim is the learning of basic laparoscopic skills, we could assume that this simulator might be a mean to complement the existing contents of basic surgical education. This perspective needs, however, to be confirmed by investigating the implementation of this simulator in an early training stage along with the study of the impact of haptic feedback on learning.

1 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 molecule present in a drug is 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 an informed consent describing their rights. With respect to 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.

Availability of data and materials

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

Competing interests

The authors declare that they have no competing interests.

Funding

Funding has been provided by the ANR (French National Research Agency) for the SAMSEI project (ANR-11-IDFI-0034) under the supervision of Pr. X. Martin.

Author's contributions

BDW and CB actively designed and realized the prototype of this simulator. RM, AL, NH supervised this work. XM and CC oversaw it. All these people contributed equally to the writing of this paper.

Acknowledgements

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

Author details

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

References

1. Roumm, A.R., Pizzi, L., Goldfarb, N.I., Cohn, H.: Minimally invasive: minimally reimbursed? an examination of six laparoscopic surgical procedures. Surgical innovation **12**, 261–87 (2005)
2. Sheetz, K.H., Norton, E.C., Birkmeyer, J.D., Dimick, J.B.: Provider experience and the comparative safety of laparoscopic and open colectomy. Health Services Research **52**(1), 56–73 (2017). doi:10.1111/1475-6773.12482. https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6773.12482
3. Spanjersberg, W.R., van Sambeeck, J.D.P., Bremers, A., Rosman, C., van Laarhoven, C.J.H.M.: Systematic review and meta-analysis for laparoscopic versus open colon surgery with or without an eras programme. Surgical endoscopy **29**, 3443–53 (2015)
4. Langlois, J., Bellemare, C., Toulouse, J., Wells, G.A.: Spatial abilities and technical skills performance in health care: a systematic review. Medical Education **49**(11), 1065–1085 (2015). doi:10.1111/medu.12786. https://onlinelibrary.wiley.com/doi/pdf/10.1111/medu.12786
5. Molinas, C.R., Binda, M.M., 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. Gynecological Surgery **14**(1), 12 (2017)
6. 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. International Journal of Urology **24**(8), 639–640 (2017). doi:10.1111/iju.13371
7. Nepomnayshy, D., Alseidi, A.A., Fitzgibbons, S.C., Stefanidis, D.: Identifying the need for and content of an advanced laparoscopic skills curriculum: results of a national survey. The American Journal of Surgery **211**(2), 421–425 (2016). doi:10.1016/j.amjsurg.2015.10.009
8. Reznick, R.K., MacRae, H.: Teaching surgical skills — changes in the wind. New England Journal of Medicine **355**(25), 2664–2669 (2006). doi:10.1056/nejmra054785
9. Schmitt, F., Mariani, A., Eyssartier, E., Granry, J.-C., Podevin, G.: Learning laparoscopic skills: Observation or practice? Journal of Laparoendoscopic & Advanced Surgical Techniques **28**(1), 89–94 (2018). doi:10.1089/lap.2017.0254
10. Badash, I., Burtt, K., Solorzano, C.A., Carey, J.N.: Innovations in surgery simulation: a review of past, current and future techniques. Annals of Translational Medicine **4**(23), 453–453 (2016). doi:10.21037/atm.2016.12.24
11. Amitai Ziv, P.R.W. Stephen D. Small: Patient safety and simulation-based medical education. Medical Teacher **22**(5), 489–495 (2000). doi:10.1080/01421590050110777
12. Dunkin, B., Adrales, G.L., Apelgren, K., Mellinger, J.D.: Surgical simulation: a current review. Surgical Endoscopy **21**(3), 357–366 (2006). doi:10.1007/s00464-006-9072-0
13. Munshi, F., Lababidi, H., Alyousef, S.: Low- versus high-fidelity simulations in teaching and assessing clinical skills. Journal of Taibah University Medical Sciences **10**(1), 12–15 (2015). doi:10.1016/j.jtumed.2015.01.008

14. Spruit, E.N., Band, G.P.H., Hamming, J.F., Ridderinkhof, K.R.: Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. Psychological Research **78**(6), 878–891 (2013). doi:10.1007/s00426-013-0525-5
15. Laguna, M.P., de Reijke, T.M., Wijkstra, H., de la Rosette, J.: Training in laparoscopic urology. Current Opinion in Urology **16**(2), 65–70 (2006). doi:10.1097/01.mou.0000193377.14694.7f
16. Gallagher, A.G., Ritter, E.M., Champion, H., Higgins, G., Fried, M.P., Moses, G., Smith, C.D., Satava, R.M.: Virtual reality simulation for the operating room. Annals of Surgery **241**(2), 364–372 (2005). doi:10.1097/01.sla.0000151982.85062.80
17. 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. Journal of Surgical Education **69**(2), 190–195 (2012). doi:10.1016/j.jsurg.2011.07.011
18. Maran, N.J., Glavin, R.J.: Low- to high-fidelity simulation - a continuum of medical education? Medical Education **37**(s1), 22–28 (2003). doi:10.1046/j.1365-2923.37.s1.9.x
19. Parmar, S., Delaney, C.P.: The role of proximate feedback in skills training. The Surgeon **9**, 26–27 (2011). doi:10.1016/j.surge.2010.11.006
20. Hammoud, M.M., Nuthalapaty, F.S., Goepfert, A.R., Casey, P.M., Emmons, S., Espey, E.L., Kaczmarczyk, J.M., Katz, N.T., Neutens, J.J., Peskin, E.G.: To the point: medical education review of the role of simulators in surgical training. American Journal of Obstetrics and Gynecology **199**(4), 338–343 (2008). doi:10.1016/j.ajog.2008.05.002
21. McKenna, D.T., Mattar, S.G.: What is wrong with the training of general surgery? Advances in Surgery **48**(1), 201–210 (2014). doi:10.1016/j.yasu.2014.05.010
22. 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. Surgical Endoscopy **31**(6), 2474–2482 (2016). doi:10.1007/s00464-016-5249-3
23. Aggarwal, R., Crochet, P., Dias, A., Misra, A., Ziprin, P., Darzi, A.: Development of a virtual reality training curriculum for laparoscopic cholecystectomy. British Journal of Surgery **96**(9), 1086–1093 (2009). doi:10.1002/bjs.6679
24. Aggarwal, R., Tully, A., Grantcharov, T., Larsen, C., Miskry, T., Farthing, A., Darzi, A.: Virtual reality simulation training can improve technical skills during laparoscopic salpingectomy for ectopic pregnancy. BJOG: An International Journal of Obstetrics & Gynaecology **113**(12), 1382–1387 (2006). doi:10.1111/j.1471-0528.2006.01148.x
25. Li, M.M., George, J.: A systematic review of low-cost laparoscopic simulators. Surgical Endoscopy **31**(1), 38–48 (2016). doi:10.1007/s00464-016-4953-3
26. Stefanidis, D., Acker, C., Heniford, B.T.: Proficiency-based laparoscopic simulator training leads to improved operating room skill that is resistant to decay. Surgical Innovation **15**(1), 69–73 (2008). doi:10.1177/1553350608316683
27. Noble, C.: The relationship between fidelity and learning in aviation training and assessment. Journal of Air Transportation World Wide **7**(3), 33–54 (2002)
28. Seymour, N.E., Gallagher, A.G., Roman, S.A., O'Brien, M.K., Bansal, V.K., Andersen, D.K., Satava, R.M.: Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Annals of surgery **236**(4), 458–464 (2002)
29. 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. Journal of Surgical Research **213**, 69–74 (2017).

doi:10.1016/j.jss.2017.02.038

30. Brinkmann, C., Fritz, M., Pankratius, U., Bahde, R., Neumann, P., Schlueter, S., Senninger, N., Rijcken, E.: Box- or virtual-reality trainer: Which tool results in better transfer of laparoscopic basic skills?—a prospective randomized trial. Journal of Surgical Education **74**(4), 724–735 (2017). doi:10.1016/j.jsurg.2016.12.009
31. Poursartip, B., LeBel, M.-E., Patel, R.V., Naish, M.D., Trejos, A.L.: Analysis of energy-based metrics for laparoscopic skills assessment. IEEE Transactions on Biomedical Engineering **65**(7), 1532–1542 (2018). doi:10.1109/tbme.2017.2706499
32. Blackhall, V.I., Cleland, J., Wilson, P., Moug, S.J., Walker, K.G.: Barriers and facilitators to deliberate practice using take-home laparoscopic simulators. Surgical Endoscopy **33**(9), 2951–2959 (2018). doi:10.1007/s00464-018-6599-9
33. Lahanas, V., Loukas, C., Georgiou, K., Lababidi, H., Al-Jaroudi, D.: Virtual reality-based assessment of basic laparoscopic skills using the leap motion controller. Surgical Endoscopy **31**(12), 5012–5023 (2017). doi:10.1007/s00464-017-5503-3
34. Leijte, E., Arts, E., Witteman, B., Jakimowicz, J., Blaauw, I.D., Botden, S.: Construct, content and face validity of the eoSim laparoscopic simulator on advanced suturing tasks. Surgical Endoscopy **33**(11), 3635–3643 (2019). doi:10.1007/s00464-018-06652-3
35. Schout, B.M.A., Hendrikx, A.J.M., Scheele, F., Bemelmans, B.L.H., Scherpbier, A.J.J.A.: Validation and implementation of surgical simulators: a critical review of present, past, and future. Surgical Endoscopy **24**(3), 536–546 (2009). doi:10.1007/s00464-009-0634-9
36. The standards for educational and psychological testing. Technical report. `https://www.apa.org/science/programs/testing/standards`
37. Cook, D.A., Beckman, T.J.: Current concepts in validity and reliability for psychometric instruments: Theory and application. The American Journal of Medicine **119**(2), 166–716616 (2006). doi:10.1016/j.amjmed.2005.10.036
38. Fried, G.M.: FLS assessment of competency using simulated laparoscopic tasks. Journal of Gastrointestinal Surgery **12**(2), 210–212 (2007). doi:10.1007/s11605-007-0355-0
39. 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. Surgical Endoscopy **20**(8), 1275–1280 (2006). doi:10.1007/s00464-005-0036-6
40. Young, J.Q., Merrienboer, J.V., Durning, S., Cate, O.T.: Cognitive load theory: Implications for medical education: AMEE guide no. 86. Medical Teacher **36**(5), 371–384 (2014). doi:10.3109/0142159x.2014.889290
41. Mayer, R.E., Moreno, R.: Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist **38**(1), 43–52 (2003). doi:10.1207/s15326985ep3801_6
42. 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 and Virtual Environments **9**(3), 236–255 (2000). doi:10.1162/105474600566772
43. Kozak, I., Banerjee, P., Luo, J., Luciano, C.: Virtual reality simulator for vitreoretinal surgery using integrated OCT data. Clinical Ophthalmology, 669 (2014). doi:10.2147/opth.s58614
44. Morasso, P.: Spatial control of arm movements. Experimental Brain Research **42**(2), 223–227 (1981)
45. POLLICK, F.E., SAPIRO, G.: Constant affine velocity predicts the 13 power law of planar motion perception and generation. Vision Research **37**(3), 347–353 (1997). doi:10.1016/s0042-6989(96)00116-2
46. Maoz, U., Berthoz, A., Flash, T.: Complex unconstrained three-dimensional hand movement and constant equi-affine speed. Journal of Neurophysiology **101**(2), 1002–1015 (2009). doi:10.1152/jn.90702.2008

47. Cifuentes, J., Boulanger, P., Pham, M.T., 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, ??? (2014). doi:10.1109/embc.2014.6943964. https://doi.org/10.1109/embc.2014.6943964
48. Carter, F.J., Schijven, M.P., Aggarwal, R., Grantcharov, T., Francis, N.K., Hanna, G.B., Jakimowicz, J.J.: Consensus guidelines for validation of virtual reality surgical simulators. Surgical Endoscopy **19**(12), 1523–1532 (2005). doi:10.1007/s00464-005-0384-2
49. Anastasi, A.: Psychological Testing, 4th Ed., pp. –750750. Macmillan, Oxford, England (1976)
50. Cronbach, L.J.: Essentials of psychological testing. Science Education **35**(4), 227–227 (1951). doi:10.1002/sce.3730350432. https://onlinelibrary.wiley.com/doi/pdf/10.1002/sce.3730350432
51. Mathis, K.L., Wiegmann, D.A.: Construct validation of a laparoscopic surgical simulator. Simulation In Healthcare: The Journal of the Society for Simulation in Healthcare **2**(3), 178–182 (2007). doi:10.1097/sih.0b013e318137aba1
52. Gallagher, A.G., Lederman, A.B., McGlade, K., Satava, R.M., Smith, C.D.: Discriminative validity of the minimally invasive surgical trainer in virtual reality (MIST-VR) using criteria levels based on expert performance. Surgical Endoscopy **18**(4), 660–665 (2004). doi:10.1007/s00464-003-8176-z
53. Vallerand, R.J.: Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française. Canadian Psychology/Psychologie canadienne **30**(4), 662–680 (1989). doi:10.1037/h0079856
54. Maithel, 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. Surgical Endoscopy **20**(1), 104–112 (2005). doi:10.1007/s00464-005-0054-4
55. Aggarwal, R., Grantcharov, T., Moorthy, K., Hance, J., Darzi, A.: A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. The American Journal of Surgery **191**(1), 128–133 (2006). doi:10.1016/j.amjsurg.2005.10.014
56. Tchantchaleishvili, V., Myers, P.O.: Hand laterality and acquired ambidexterity in surgical training. Annals of Surgery **264**(6), 18–19 (2016). doi:10.1097/sla.0000000000001951
57. Anna, S., Gregory, A., Margaret, M., Marc, S.: Ambidexterity in laparoscopic surgical skills training. Studies in Health Technology and Informatics **184**(Medicine Meets Virtual Reality 20), 412–416 (2013). doi:10.3233/978-1-61499-209-7-412
58. Dayan, E., Cohen, L.G.: Neuroplasticity subserving motor skill learning. Neuron **72**(3), 443–454 (2011). doi:10.1016/j.neuron.2011.10.008
59. Islam, G., Kahol, K., Li, B., Smith, M., Patel, V.L.: Affordable, web-based surgical skill training and evaluation tool. Journal of Biomedical Informatics **59**, 102–114 (2016). doi:10.1016/j.jbi.2015.11.002
60. Ahmed, N., Devitt, K.S., Keshet, I., Spicer, J., Imrie, K., Feldman, L., Cools-Lartigue, J., Kayssi, A., Lipsman, N., Elmi, M., Kulkarni, A.V., Parshuram, C., Mainprize, T., Warren, R.J., Fata, P., Gorman, M.S., Feinberg, S., Rutka, J.: A systematic review of the effects of resident duty hour restrictions in surgery. Annals of Surgery **259**(6), 1041–1053 (2014). doi:10.1097/sla.0000000000000595

Figures

Figure 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.

Figure 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 5 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. Mean values (SD) for time, collisions and standard deviation means for affine velocity in exercise 1 for experts, intermediates (INT), inexperienced interns (I. Interns) and novices. *=p$<$0.05; **=p¡0.01; ***p$<$0.001

Figure 3 Mean values (SD) for time, collisions and standard deviation means for affine velocity in exercise 1. For experts, intermediates (INT), inexperienced interns (I. Interns) and novices. *=p$<$0.05; **=p$<$0.01; ***p$<$0.001

Figure 4 Mean values (SD) for time, collisions and standard deviation means for affine velocity in exercise 2. For experts, intermediates (INT), inexperienced interns (I. Interns) and novices. *=p$<$0.05; **=p$<$0.01; ***p$<$0.001

Tables

Table 1 Participants' demographic and specialty.

Group	Total Number (Male:Female)	Mean age (range)	Paediatric speciality	Gynaecology speciality	Urology speciality	Digestive speciality	Visceral speciality	Orthopaedic speciality	General
Experts	17 (13:4)	41.5 (62-30)	1	1	8	4	2	1	0
Intermediates	12 (11 :1)	28 (32-25)	0	0	10	0	2	0	0
Inexperienced Interns	28 (20 :8)	25.3 (40-23)	0	0	0	0	0	0	28
Novices	20 (7:13)	31.5 (53-18)	0	0	0	0	0	0	0