Assessment of technical skills based on learning curve analyses in laparoscopic surgery training

Hardon, Sem F.; van Gastel, Leonie A.; Horeman, Tim; Daams, Freek

DOI
10.1016/j.surg.2021.04.024

Publication date
2021

Document Version
Final published version

Published in
Surgery (United States)

Citation (APA)
Hardon, S. F., van Gastel, L. A., Horeman, T., & Daams, F. (2021). Assessment of technical skills based on learning curve analyses in laparoscopic surgery training. *Surgery (United States)*, *170*(3), 831-840. https://doi.org/10.1016/j.surg.2021.04.024

Surgery 170 (2021) 831–840

Contents lists available at ScienceDirect

Surgery

journal homepage: www.elsevier.com/locate/surg

Minimally Invasive Surgery/Endoscopy

Assessment of technical skills based on learning curve analyses in laparoscopic surgery training

Sem F. Hardon, MD[a,b,*], Leonie A. van Gastel, MD[a], Tim Horeman, PhD[b,1], Freek Daams, MD, PhD[a,1]

[a] *Department of Surgery, Amsterdam UMC–VU University Medical Center, Amsterdam, the Netherlands*
[b] *Department of Biomechanical Engineering, Delft University of Technology, Delft, The Netherlands*

ARTICLE INFO

Article history:
Accepted 25 April 2021
Available online 2 June 2021

ABSTRACT

Background: Objective force- and motion-based assessment is currently lacking in laparoscopic skills curricula. This study aimed to evaluate the added value of parameter-based assessment and feedback during training.

Methods: Laparoscopy-naïve surgical residents that took part in a 3-week skills training curriculum were included. A box trainer equipped with the ForceSense system was used for assessment of tissue manipulation- (MaxForce) and instrument-handling skills (Path length and Time). Learning curves were established using linear regression tests. Pre- and post-course comparisons indicated the overall progression and were compared to predefined proficiency levels. A post-course survey was carried out to assess face validity.

Results: In total, 4,268 trials, executed by 24 residents, were successfully assessed. Median (interquartile range) MaxForce outcomes improved from 2.7 Newton (interquartile range 1.9–3.8) to 1.8 Newton (interquartile range 1.2–2.4) between pre- and post-course assessment ($P \leq .009$). Instrument *Path length* improved from 7,102.2 mm (interquartile range 5,255.2–9,025.9) to 3,545.3 mm (interquartile range 2,842.9–4,563.2) ($P \leq .001$). Time to execute the task improved from 159.8 seconds (interquartile range 119.8–219.0) to 60.7 seconds (interquartile range 46.0–79.5) ($P \leq .001$). The learning curves revealed during what training phase the proficiency benchmarks were reached for each trainee. In the survey outcomes, trainees indicated that this curriculum should be part of a surgical residency program (mean visual analog scale score of 9.2 ± 0.9 standard deviation).

Conclusion: Force-, motion-, and time-parameters can be objectively measured during basic laparoscopic skills curricula and do indicate progression of skills over time. The ForceSense parameters enable curricula to be designed for specific proficiency-based training goals and offer the possibility for objective classification of the levels of expertise.

Introduction

Patient safety in minimally invasive surgery (MIS) must be ensured by adequate training and assessment of technical skills.[1–3] The complexity of this technique, compared to open surgery, has resulted in a more shallow learning curve (LC) with relatively little improvement after each case or trial.[4] This is explained by the loss of a degree of freedom and the loss of three-dimensional vision.[5–7] Moreover, the use of relatively long instruments resulted in the fulcrum effect, amplified tremor, and the loss of haptic feedback.[8] Over the past few decades, it has become clear that advancements in technology and expertise in training directly translated into enhanced patient safety.[9] Therefore, technical skills should preferably be mastered in advance of performing MIS in the operating room (OR). Moreover, mastery of this aspect of surgical competence should be objectively assessed.[10]

Recently, much effort was devoted to the development of performance parameters that reflect efficient and safe tissue interaction. For example, force exertion parameters (eg, MaxForce)

The preliminary result of this paper were awarded the "Gerhard Buess Technology Award" at the 25th International Congress of the European Association for Endoscopic Surgery.

Tim Horeman and Freek Daams are senior authors who contributed equally to this paper.

* Reprint requests: Sem F. Hardon, MD. Amsterdam UMC - VU Universirty Medical Center, Department of Surgery, Room ZH 7F 005, De Boelelaan 1117, 1081 HV, Amsterdam, the Netherlands.

E-mail address: s.hardon@amsterdamumc.nl (S.F. Hardon);

Twitter: @sem_hardon, @freekdaams

[1] Shared senior authorship.

https://doi.org/10.1016/j.surg.2021.04.024
0039-6060/

indicate if tissue damage occurs during manipulation, and instrument motion parameters, such as “Path length” and “Volume of motion,” indicate any risk of unintended interaction with tissue out of the camera’s scope.[11–17] Despite the development and validation of tools for objective assessment of technical skills acquisition over the past decades, structural implementation and adoption into surgical curricula are still lacking. These systems enable the earlier mentioned parameter-based assessment of gentle tissue manipulation (force-based parameters) and instrument handling skills (motion-based parameters and time). Trainers utilizing these tools can be facilitated and objectively informed about whether their trainee has reached the desired level of technical skills and, thus, whether the LC for the above-mentioned skills is passed sufficiently before operating on real patients.

Patient safety may be improved if novice laparoscopic surgeons perform their first operation on a real patient after reaching the plateau phase for technical skills in a simulation environment.[18] However, accumulating evidence reveals the individual variability in the gradient of the LC.[19] In contrast to trainees possessing innate technical ability, previous research showed that the vast amount of trainees appeared to be moderate or low performers (71–90%).[20] This group struggled to achieve progression with every trial, and some never met the predefined level of competence (ie, the predefined level of proficiency) at the end of the training.

In this present study, we aimed to map individual differences in LC by using force- and motion-based assessments. We sought to determine whether parameter outcomes can be utilized to assess the differences in LC among individual trainees in order to optimize feedback during training. We hypothesized that the assessment tool would be able to discriminate between trainees and that a wide range of technical skill sets during training would be revealed. Moreover, we aimed to provide an indication on whether the plateau phase has been reached to indicate if the trainees have acquired the requisite level of proficiency before operating on real patients.

Methods

This prospective cohort study was conducted between April 2018 and August 2019. A novel designed and validated training program for fundamental laparoscopic skills (FLS) training was implemented into clinical practice.

Participants

In this multidisciplinary setup, first-year general surgical residents with no or limited prior laparoscopic experience were included as a trainee (Supplementary Table S1). All participants were affiliated with the Amsterdam University Medical Center (Amsterdam UMC) and received their training at either the Departments of General Surgery, Urology, Plastic and Reconstructive Surgery, or Orthopedic Surgery. Trainees were enrolled in this national training program as part of the first 2 years of their regular general surgery residency program in the Netherlands.

Systems and hardware

A portable box trainer (Supplementary Figure S1) measuring 45 × 30 × 25 cm (LAPSTAR training system, Camtronics B.V., Son en Breugel, the Netherlands) was used for all experiments. The box was equipped with the ForceSense system (MediShield B.V., Delft, the Netherlands) for objective tracking, monitoring, and assessment of tissue interaction forces and instrument handling metrics.[16,21] This system was connected to a user interface (Windows Surface tablet, Microsoft, Redmond, WA) with ForceSense.net client software to track performances and to facilitate direct online data logging on the website that was in concordance with local privacy regulations.

The task set (Supplementary Figure S2) consisted of 4 previously validated task inserts (3Dmed, Franklin, OH) commonly used to train FLS.[18,22,23] A fifth task (Medishield, Delft, the Netherlands) was developed and validated at the Department of Biomechanical Engineering at Delft University of Technology, specifically for training tissue manipulation skills and bimanual dexterity.[16,24] Moreover, the potential value of this task was previously demonstrated in a pilot study.[18] The sixth task is known as the “Pattern Cut Tasks,” which is commonly known to be part of the FLS curriculum by the American Board of Surgery.[23,25] Two curved Maryland dissection forceps and 1 laparoscopic scissors (Aesculap, Bbraun, Melsungen, Germany) were used.

Training

At first, participants were instructed to perform an intracorporeal suturing task, which was not part of the course, to familiarize themselves with the instruments and equipment without influencing the LC. After successful completion, participants performed the 6 training tasks per-protocol (Supplementary Figure S3) in a fixed order as a baseline assessment. Subsequently, the box trainer was taken home for 3 consecutive weeks, during which trainees were instructed to autonomously train a minimum of 4 sessions per week, with a minimum of 15 minutes for each training session. All participants were assigned to train the tasks in the same fixed order and to move only to the next task when the preset training goal was reached. This training goal was established during a previous validation study.[18] The mean parameter outcomes of 7 laparoscopic surgeons (number of advanced procedures >50) were considered the proficiency benchmark for this course. After passing the predefined proficiency level for the 6 tasks, trainees were allowed to continue their mastery of skills beyond this level, as long as their intrinsic motivation was not depleted. The 6 tasks were performed consecutively again as a post-course assessment to determine the effectiveness of this course in terms of progression at the end of the 3-week course. As all parameter data and videos were stored at the ForceSense.NET database, the progress of trainees was monitored and skills levels were registered to ensure that each participant executed the training properly.

Measurements and feedback

The ForceSense measurement system measures and records a total of 17 parameters for each trial. The 5 different tissue handling parameters are based on 3D measurement of force appliance and manipulation of tissue. The 12 motion analyses parameters (MAP) (eg, path length, insertion depth, and motion volume) are based on 3D measurements of tip position and motion analysis information of both instruments. Based on previous validation studies executed with the proposed training tasks, the 7 most discriminating and relevant parameters regarding the assessment of tissue trauma were analyzed.[16,18,26] A detailed overview and description of these parameters are provided in Table I. Multi-parametric objective feedback was provided by the user interface directly after each trial to facilitate and contribute to the concept of deliberate practice.[27]

Table I
Description of parameters

Parameter	Unit	Description
Time	Seconds	Time measured from the beginning of the task until task completion.
Path length	Millimeters	Total distance travelled by the tip of the right and left instrument together during the task.
Mean force NZ (non-zero)	Newton	Mean absolute force exerted by the instruments on the task platform during periods when force is not zero.
Maximal force	Newton	The highest absolute force as exerted by the instruments on the task platform during the task.
Maximal impulse	Newton seconds	The largest product of force and the duration that the force was exerted, before force returned to zero (the area under a force peak if force is presented in time).
SD Force	Newton	The standard deviation of the absolute force.
Force volume	Cubic Newton (N^3)	The volume of an ellipsoid fitted around the standard deviations of the force along 3 principal components. High force volume indicates fast increasing and decreasing forces in different direction.

The trainees received feedback on the number of force penalties (number of times exceedance of the threshold for maximum force exertion), total path length, and time directly after each executed trial. In addition, to enable instant tracking of their progression, trainees received access to the online data platform for an overview of their results and visualization of the LC for each parameter.

Face validity

All participants were asked to fill out a questionnaire after completion of the 3-week curriculum to obtain information about the protocol content and their general impression of the box trainer and training task. Moreover, they were asked their general opinion with regard to simulation training. Responses are presented on a visual analog scale (VAS).

LC evaluation and analysis

Individual LC was visualized to receive information on the character, slope, and gradient of the LC for each participant. Graphs were created using the relationship between the number of repetitions and the moving averages of parameter outcomes for each task. Statistical methods can be used to analyze LC. Instead of group-based comparison, as is usual in construct validation studies,[28] we depicted line graphs showing the relationship between the number of repetitions and the 7 parameter outcomes created after the data was smoothened with the centered moving average procedure using a span of 3. To evaluate tasks, specific characteristics, and to define recommendations for implementation of these tasks in skills curricula, we additionally analyzed and plotted the mean outcomes of this cohort. The minimum amount of trials executed by all participants was 10. The mean of these outcomes was used to analyze the group LC and to evaluate the differences in gradient between parameters. Least squares regression analysis was used for curve fitting: fitting a model to the graph, based on a mathematical equation.[29] An inverse model ($Y = a + b/X$) was used between every succeeding trial to estimate the plateau phase of the LC,[30] in which Y represents the parameter outcomes and X represents the number of repetitions. The plateau was defined as the theoretical best outcome, achieved when the number of repetitions is infinite ($Y = a$ when $X = \infty$). To ensure that the trainee's calculated plateau actually indicates the sufficient level of skills to perform surgery on real patients, we compared the outcomes with the earlier mentioned predefined level of proficiency. IBM SPSS Statistics for Windows, version 26 (IBM Corp, Armonk, NY) was used to calculate the values for a and b for all parameters, for each participant. Subsequently, it was estimated how many repetitions the participants needed to reach within 20% of their theoretical best achievable score, calculated as $X = 5 * b/a$ ($X = 5 * b/a$ when $Y = 1{,}2a$). In addition to this, the number of repetitions needed to reach the predefined proficiency level was calculated.

Additional statistical analysis

To assess the improvement of the residents' technical skill sets, pre- and post-course parameters of the 6 training tasks were analyzed and compared. All performances that were not completed successfully (ie, terminated prematurely due to unforeseen circumstances; a bot that is boiling over, ringing doorbell, etc.) were excluded from analyses. Since the Kolmogorov-Smirnov test indicated that data were not normally distributed, the non-parametric Wilcoxon signed-rank test (for paired samples) was used. Data were again analyzed using IBM SPSS Statistics for Windows, version 26 (IBM Corp, Armonk, NY).

Results

In total, 4,268 trials were executed by 24 participants (16 female; mean age 29 (± 1.9 standard deviation [SD]); 23 right-handed) and included for analyses. In total, 278 performances were terminated prematurely due to unforeseen circumstances and excluded from analyses. At the baseline evaluation, 19 trainees had no prior experience with laparoscopic box training and 4 trainees had trained before; 1–5 times ($n = 3$) or >20 times ($n = 1$). The experience with laparoscopic surgery as chief operator during basic laparoscopic procedures (appendectomy, cholecystectomy, or hernia repair) was limited to none ($n = 11$), 1–10 times ($n = 10$), or 11–50 times ($n = 3$).

LC analyses

Similar trends in LC are seen for instrument handling parameters *Time* and *Path Length* (Figs 1, *A*, *B* and 2, *A*, *B*) and Supplementary Figure S4), with rapid improvement over the first trials and then gradually levels out until the plateau phase is reached. The LC representing tissue manipulation skills (Figs 1, *C* and 2, *C*) show more abrupt, step-like curves, with alternating improvement and decrease in skills after each trial and no association with the MAP. Table II shows the comparison of the median number of trials to reach the plateau phase (column A) with the median number of trials to reach the level of proficiency (column B) for each task. For all tasks, except task 4, proficiency levels for maximal force have been reached before the LC was leveled out into the plateau phase. The same accounts for maximal impulse (5 out of 6 tasks), force volume (5 out of 6 tasks), time (4 out of 6 tasks), and path length (3 out of 6 tasks).

Overall progression of skills

All participants succeeded to acquire the predefined levels of proficiency (Table III). Considerable variation in the initial skills set between trainees was present for all parameters at the pre-course assessment. A reduction in force-based and MAP outcomes, with

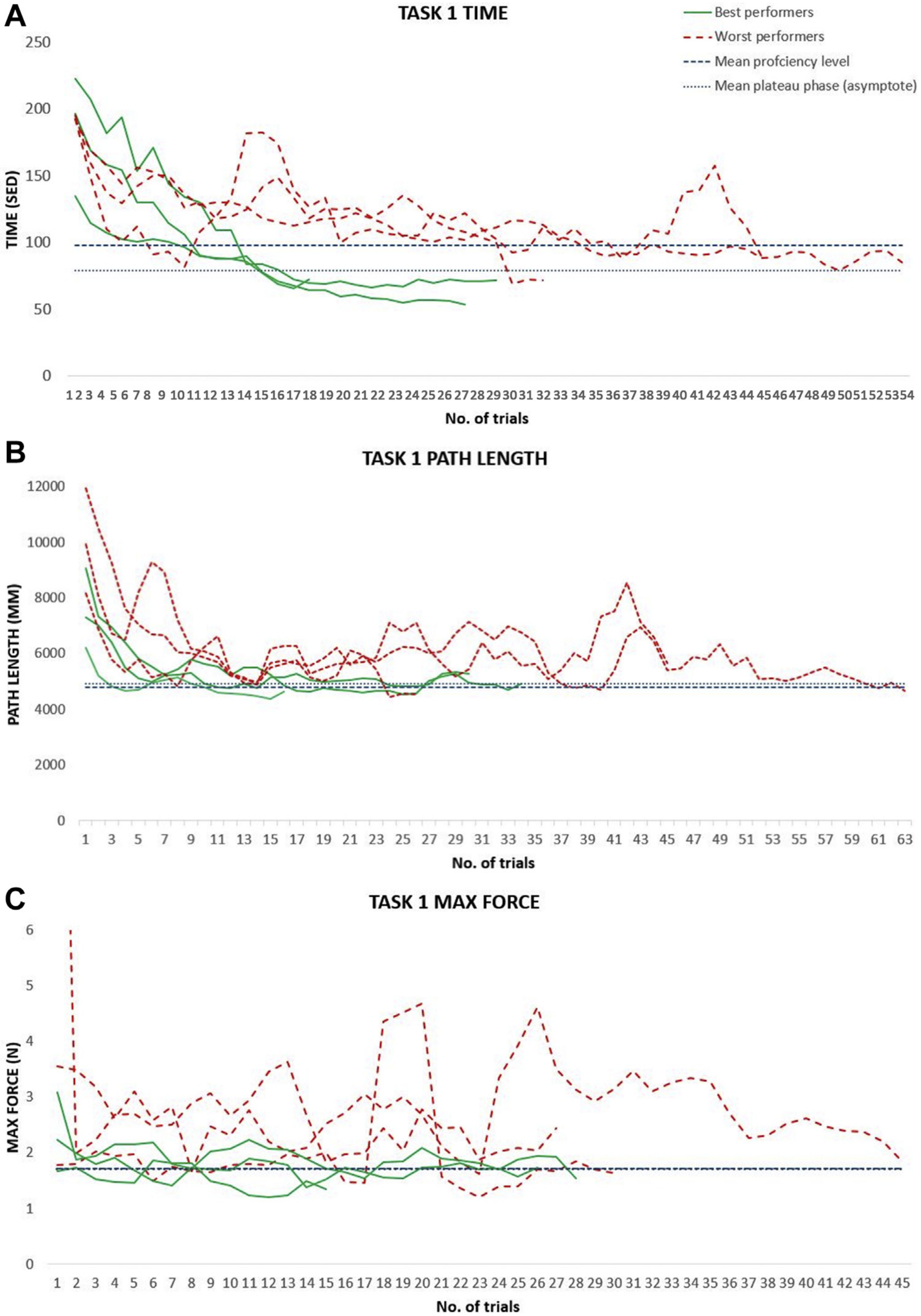

Fig 1. (A–C) Individual learning curve pattern among best and worst performers for task 1 "Post and Sleeve" as one of the most representative and discriminating tasks, based on the ability to reach both the proficiency level and plateau phase of learning curve

a convergence of (interquartile) ranges is observed in the post-course assessment outcomes. In 5 out of 6 tasks, we observed that (median and interquartile range [IQR]) maximal force parameter outcomes improved from 2.7 Newton (IQR 1.9–3.8) to 1.8 Newton (IQR 1.2–24) between pre- and post-course assessment ($P \leq .009$). Mean overall progression in technical skills was 31.7% (SD ± 11.6%). Significant reductions of maximum exerted forces were observed in 5 tasks ($P \leq 0.009$), and a reduction of maximal impulse was observed for all tasks ($P \leq .015$). The largest improvement of tissue handling outcomes was seen in task 3 (Flap task), with maximal force parameters having improved up to 47% and maximal impulse parameter even improved by 89%. This is consistent with other force-based parameter outcomes for this task.

MAP significantly improved in all tasks; exerted instrument path length improved from 7,102.2 mm (IQR 5,255.2–9,025.9) to 3,545.3 mm (IQR 2,842.9–4,563.2) ($P \leq .001$). Time to execute the task improved from 159.8 seconds (IQR 119.8–219.0) to 60.7 seconds (IQR 46.0–79.5) ($P \leq .001$). Mean overall progression in instrument handling was 49.8% (SD ± 9.7%) for path length and 60.5% (SD ± 10.5%) for time. Most progression of efficiency (ie, time) was observed in tasks 2 (Loops and Wire; 71%).

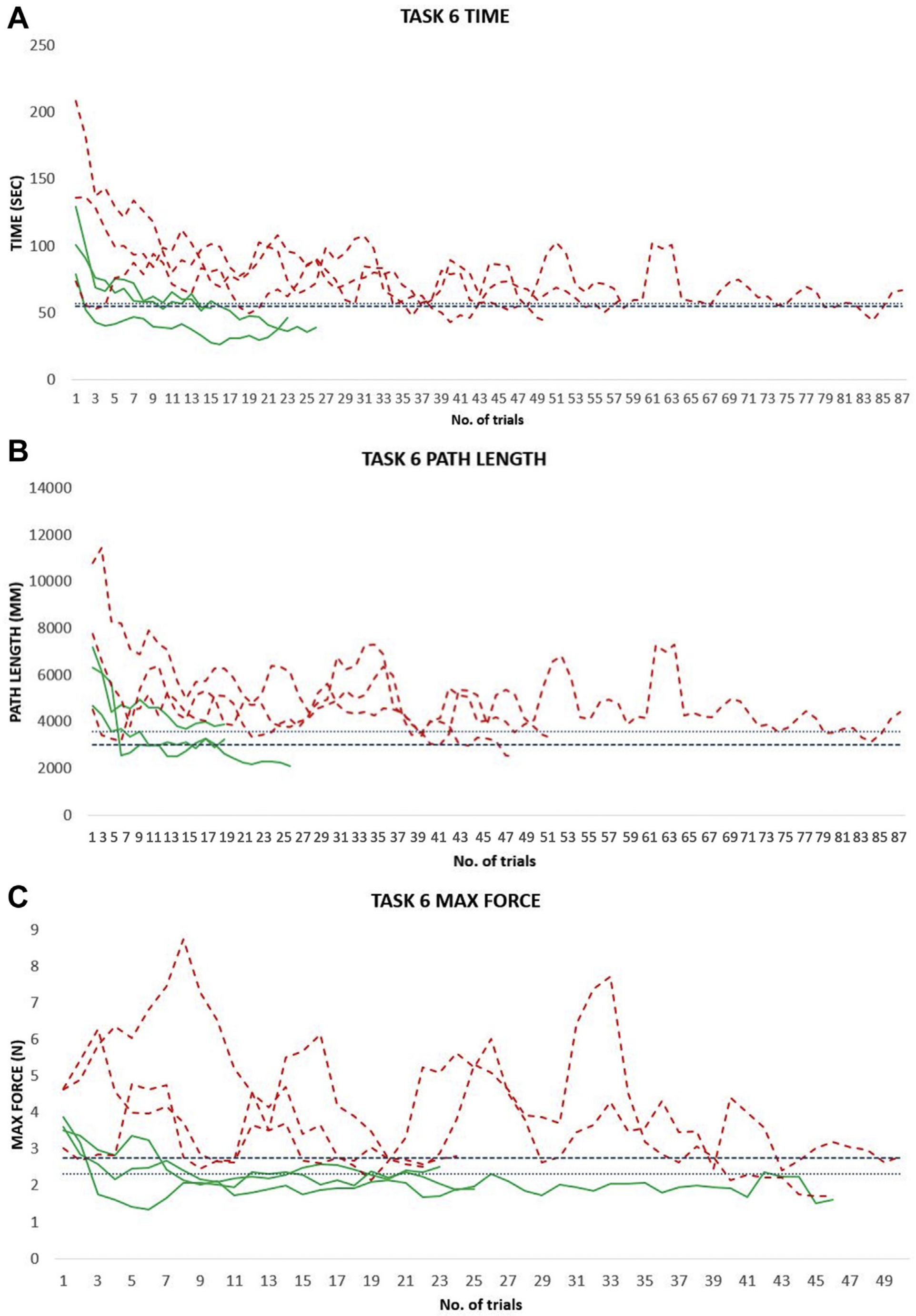

Fig 2. (A–C) Individual learning curve pattern among best and worst performers for task 6 "Zig-zag loop" as one of the most representative and discriminating tasks, based on the ability to reach both the proficiency level and plateau phase of learning curve

Figure 3 displays the relationship between the number of performances and the mean moving average of the cohort for the first 10 repetitions of each task for all 7 parameters separately. Force-based parameters tend to show limited improvement over time for most tasks (Fig 3, *C*–*F*), except for task 1 (Post and sleeve; Fig 3, *D*). Also, higher forces were applied during the performance of task 2 (Loops and Wire) and task 6 (Zig-zag loop), compared to the other 4 tasks. For instrument handling parameters (MAP), similar LC with rapid improvement during the first repetitions are visible for all tasks (Fig 3, *A* and *B*).

Face validity

All participants completed the questionnaire ($N = 24$). The overall assessment of the protocol content is considered to be "good," with a mean VAS score of 8.4 ± 1.7 SD (Table IV). The self-rating of participants indicated that their technical skills (8.09 ± 1.05 SD) and self-confidence (7.70 ± 1.23 SD) were improved during this curriculum. Moreover, trainees indicated that this curriculum should be part of a surgical residency program (9.2 ± 0.9 SD).

Table II
Overall progression of skills: median and interquartile range (IQR) of pre- and post-course parameters (Wilcoxon signed-rank test)

Parameter	Task	n	Pre-course assessment	Post-course assessment	% progression	P value
Time, s (IQR)	1	24	196.70 (167.29–227.19)	73.69 (66.57–88.18)	62	<.001
	2	24	193.20 (133.42–250.62)	55.64 (48.17–66.88)	71	<.001
	3	24	114.59 (84.15–169.20)	39.04 (30.50–51.43)	66	<.001
	4	19	186.15 (116.55–242.00)	70.17 (57.83–83.76)	62	<.001
	5	24	144.16 (119.77–174.82)	90.48 (63.88–100.26)	38	<.001
	6	24	132.85 (118.08–160.92)	47.72 (37.99–60.86)	64	<.001
Path length, mm (IQR)	1	24	8,552.84 (7,601.10–11,695.85)	4,819.95 (4,529.74–5,319.88)	44	<.001
	2	24	6,956.16 (5,502.84–9,897.59)	3,363.81 (2,887.18–3,618.86)	52	<.001
	3	24	5407.66 (2,562.37–7397.24)	1,906.81 (1,484.17–2,237.06)	65	<.001
	4	19	7101.27 (5,762.47–9125.17)	4,003.94 (3,745.46–4,721.11)	44	<.001
	5	24	5719.83 (4979.47–8,136.57)	3,673.46 (3,073.66–4,339.74)	36	.001
	6	24	7,018.99 (5,238.52–8,653.37)	2,975.00 (2,638.35–4,064.29)	58	<.001
Mean force NZ, N (IQR)	1	24	0.54 (0.44–0.63)	0.51 (0.47–0.56)	–	.063
	2	24	0.85 (0.68–0.94)	0.69 (0.61–0.79)	19	.001
	3	24	0.51 (0.46–0.57)	0.40 (0.35–0.43)	22	<.001
	4	19	0.53 (0.44–0.59)	0.47 (0.44–0.53)	–	.119
	5	24	0.60 (0.52–0.71)	0.48 (0.42–0.72)	–	.255
	6	24	0.83 (0.74–1.00)	0.65 (0.60–0.73)	22	<.001
Maximal force, N (IQR)	1	24	2.50 (2.16–2.90)	1.71 (1.49–2.01)	32	.002
	2	24	3.42 (2.73–4.13)	2.13 (1.94–2.72)	38	.001
	3	24	1.77 (1.50–2.39)	0.93 (0.82–1.13)	47	<.001
	4	19	2.00 (1.62–2.85)	1.66 (1.18–2.03)	17	.009
	5	24	1.99 (1.70–2.68)	1.49 (1.12–2.41)	–	.215
	6	24	4.22 (3.76–5.26)	2.55 (2.24–3.05)	40	<.001
Maximal impulse, Ns (IQR)	1	24	4.22 (2.10–6.70)	1.39 (1,12–2,54)	67	.001
	2	24	20.11 (12.01–25.17)	5.83 (4.01–7.96)	71	<.001
	3	24	7.12 (4.09–10.16)	0.80 (0.50–1.47)	89	<.001
	4	19	5.75 (3.96–13.67)	2.37 (1.73–3.33)	59	.010
	5	24	10.92 (7.29–17.61)	5.14 (2.62–16.12)	52	.015
	6	24	12.28 (9.18–17.64)	4.65 (3.22–6.50)	62	<.001
SD force, N (IQR)	1	24	0.28 (0.20–0.35)	0.24 (0.19–0.28)	14	.018
	2	24	0.57 (0.42–0.63)	0.40 (0.34–0.51)	30	.003
	3	24	0.27 (0.22–0.29)	0.13 (0.11–0.17)	52	<.001
	4	19	0.28 (0.21–0.34)	0.24 (0.19–0.27)	–	.061
	5	24	0.34 (0.25–0.41)	0.22 (0.17–0.40)	–	.173
	6	24	0.58 (0.51–0.74)	0.42 (0.35–0.47)	28	<.001
Force volume, N^3 (IQR)	1	24	0.11 (0.05–0.17)	0.07 (0.05–0.09)	36	.006
	2	24	0.24 (0.15–0.46)	0.12 (0.08–0.15)	50	<.001
	3	24	0.04 (0.02–0.06)	0.01 (0.01–0.02)	75	<.001
	4	19	0.05 (0.03–0.11)	0.05 (0.03–0.07)	44	.014*
	5	24	0.09 (0.05–0.10)	0.05 (0.03–0.09)	–	.162
	6	24	0.32 (0.19–0.59)	0.13 (0.09–0.20)	59	<.001

* *P* value was calculated using "order number" instead of the median.

Table III
Individual number of trials needed to reach the learning curve plateau (A*) and number of repetitions needed to pass the level of proficiency (B†)

	Max force		Max impulse		Mean force		Standard deviation of force		Force volume		Path length		Time	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B
Task														
1	4 (333)	4 (14)	21 (46)	6 (18)	2 (4)	5 (26)	4 (51)	5 (18)	6 (103)	5 (21)	4 (9)	11 (24)	9 (14)	7 (23)
2	6 (4)	2 (16)	17 (18)	2 (5)	2 (2)	2 (26)	3 (4)	2 (11)	10 (8)	1 (5)	7 (15)	10 (46)	12 (24)	9 (38)
3	6 (10)	2 (8)	30 (139)	3 (14)	2 (2)	2 (12)	6 (7)	2 (12)	13 (33)	2 (4)	9 (18)	6 (34)	10 (33)	10 (29)
4	4 (7)	5 (17)	25 (85)	3 (16)	2 (4)	3 (21)	3 (6)	3 (17)	6 (32)	1 (28)	5 (8)	4 (19)	12 (33)	3 (13)
5	7 (4)		12 (9)		3 (2)		6 (0)		22 (28)		7 (10)		5 (9)	
6	5 (11)	4 (14)	12 (36)	3 (6)	2 (4)	2 (14)	4 (10)	3 (14)	12 (259)	2 (13)	6 (12)	3 (9)	9 (12)	8 (36)

* Median (range) estimates of the individual learning curve plateau, defined as 'a' in '$Y = a + b/X$'.
† Median (range) of repetitions needed to reach the predefined level of proficiency (ie, mean of expert surgeons).

Discussion

This study showed that objective assessment of LC in laparoscopic training is possible for novice surgical trainees. The ForceSense system showed progress over time and could indicate the level of skills compared to a predefined proficiency level. Moreover, repetitive measurements provided detailed insight into the slope of the individual LC (Figs 1 and 2). Instant feedback after each trial enabled the introduction of alternative training in an early stage of the LC if a trainee, even with continued practice, is not capable of achieving technical proficiency (Figs 1 and 2, dashed red line).

Other than linking to trainees' skills, unexpected irregular LC fluctuations provide information about instrument difficulties or task execution inconsistencies that can help to assess the quality of used instruments and realism of used training tasks. The intuitiveness to detect the right strategy for optimal execution of the task depends on task-related factors like stiffness, smoothness,

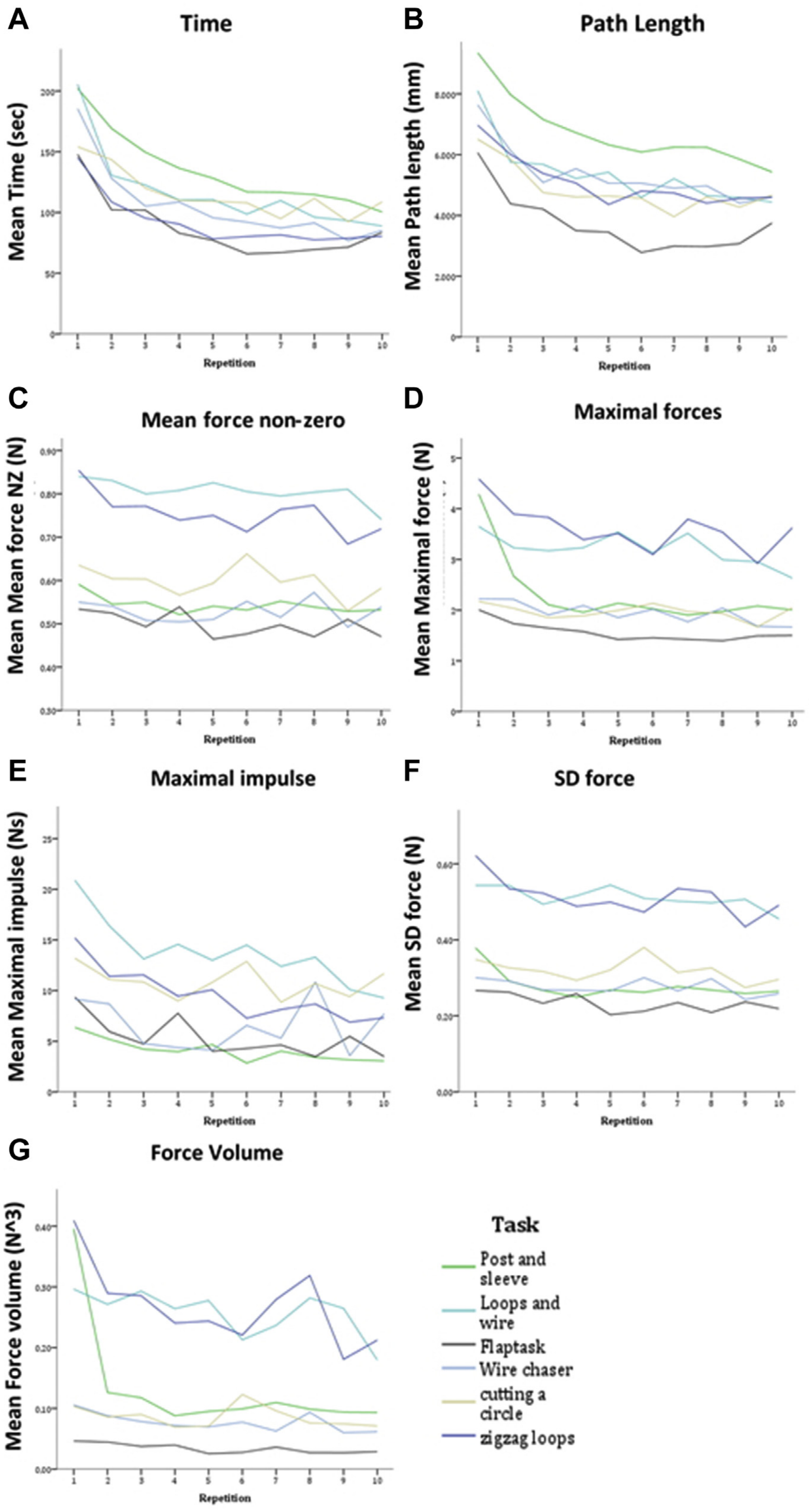

Fig 3. (A–G) Mean outcomes of the initial learning curves to evaluate differences in parameter behavior ($N = 24$).

elasticity, and friction of the artificial tissue. For example, smooth surfaces with high elasticity will result in less peak force than stiff materials with high resistance and friction. Therefore, besides the innate abilities of the trainees, the training goals and the given time span of the curriculum should be aligned with the characteristics and the complexity of the tasks.

Table IV
Face validity—post-course survey outcomes

Statement	Visual Analog Scale 0–10 cm Presented as mean (±SD)
Protocol content	
The box is valuable (useful) for laparoscopic training.	8.90 (±1.09)
How suitable are the tasks for acquisition of basic laparoscopic skills?	8.58 (±1.18)
The ForceSense metrics provide sufficient feedback on my laparoscopic skills.	6.79 (±1.94)
Training at home to develop FLS should be mandatory.	7.78 (±2.03)
The curriculum should be part of the regular surgical resident training.	9.22 (±0.88)
Timespan of the curriculum is adequate.	7.97 (±2.04)
Box trainer and tasks	
The box is easy to set up at home	8.27 (±1.65)
How well do the tasks test your laparoscopic skills	7.28 (±1.52)
How useful are the tasks for laparoscopic training?	7.94 (±1.52)
General opinion with regard to training	
I have other surgical interests/ambitions in surgery than MIS.	4.24 (±2.80)
Training should be mandatory before practicing laparoscopy at the OR.	5.45 (±2.78)
I am motivated to train at home.	7.96 (±1.55)
I prefer training in a skills lab.	5.07 (±2.47)
My skills are improved.	8.09 (±1.05)
My self-confidence considering performing laparoscopic surgery is improved.	7.70 (±1.23)

FLS, fundamental laparoscopic skills; *MIS,* minimally invasive surgery; *OR,* operating room; *SD,* standard deviation.

As stated by Hopper et al (2007), the ideal surgical LC tends to show a steep curve at first, which then fades out to a more gradual LC as the plateau phase (ie, asymptote) is approached. It is from this point that technical skill sets have been sufficiently established to independently perform surgery in a safe manner.[31,32] Based on our previous results,[16,18,33] we conclude that the decrease of the applied forces toward an expert-based proficiency level in FLS learning will result in the reduction of unintentional tissue trauma. Moreover, it is known that off-camera movements pose a potential risk for injuries, such as inadvertent instrumental perforation or capacitive coupling when using coagulation. These injuries may be avoided if laparoscopy novices optimize their instrument handling skills in order to reduce the total instrument path length.

However, the present results indicate that some trainees stop and move on to the next task, while still in the part of training where progression is made with each trial. If the LC is still rapidly descending, although the proficiency level has been passed, trainees still experience benefits from performing additional trials. If the plateau phase is reached before approaching the proficiency level, it will not be beneficial to execute additional trials, and one should investigate why the trainee was not able to reach the threshold values on a deeper psychological or motor skills level. We would like to emphasize the variability in force-based assessment outcomes (Figs 1, *C* and 2, *C*), with respect to their potential value as a predictor for tissue trauma.[16,29,34] Trainers can now be objectively informed to assess whether the personal plateau phase has been approached sufficiently, and, ultimately, to evaluate if the trainee is considered to be prepared to perform laparoscopy in the OR.

In line with previous results, significant improvement of tissue manipulation (32%) and instrument handling (≥51%) skills was observed (Table III).[18,24,35] Improvements of instrument handling parameters (ie, MAP and time) had a gradual and predictable nature. In contrast, this study reveals that LC for gentle handling of tissues (ie, force-based parameters)[34,36] resulted in dissimilar slopes among trainees, with large offset from the 0-line and a much more unpredictable character (Figs 1, *C* and 2, *C*, and Supplementary Figure S5). This implies that training in these skills may require different strategies for feedback and assessment to improve. To date, no studies have reported on using objective force-based metrics for assessment of tissue handling during training in a skills curriculum that is implemented into the surgical residency training program. By doing so, we were able to grade the shape of the LC and its discrepancies with respect to the ideal LC as described previously.[31]

Previous studies regarding the training of FLS focused on describing or validating one aspect of a curriculum (ie, the protocol, the task set, the type of assessment, or its reliability).[37–40] If the focus was on the type of assessment and its validity evidence, the vast amount of data was collected in a research setup (eg, laboratory or skills center), which was developed specifically for the purpose of executing the studies. If assessment during training was enhanced with metrics provided by objective measurement tools, previous studies[3,35,41–46] mainly reported on MAP as a measure for instrument handling skills. To date, no objective assessment of tissue manipulation skills, although of added value in experimental setups,[16,24,47–51] has been implemented into residency training programs. Since instrument handling skills (ie, MAP and Time) and tissue manipulation (ie, MaxForce) are not correlated,[16] a separate assessment of both skills is essential.[52]

Although this study was conducted before the present coronavirus disease 2019 pandemic, these outcomes have become even more relevant in light of the current regulations. Even today, while still facing several of such regulations, this course is running and residents are being trained. In order to adhere to coronavirus precautions, the system is cleaned with alcohol wipes and checked between participants, costing around 30 minutes per system. Trainers can access the data and video of all performances to ensure adequate practice. Especially in the absence of hands-on training and assessment in a skills center, these can be used for telementoring during video-conferenced group sessions with remote feedback and assessment of technical skills. However, the ForceSense.NET software should be modified to allow the streaming of introductory videos in case of a pandemic that restricts direct contact.

During this curriculum, conducted with 2 prototype box trainers, we initiated the development of a new fully integrated box trainer (Supplementary Figure S6). The ForceSense systems used in this study were purchased from the university MedTech starter MediShield BV and typically cost between 12k–20k euro, depending on the components, the type of box trainer, and license. In case task components are missing or parts are broken, it is advised to have some spare replacements on location. During this study, hardware failure occurred 2 times, and replacement of a Trendo (trocar) was needed once. The ForceSense system was created based on the ForMoST system developed at the Delft University of Technology by Horeman et al[53] and is widely available in Europe, Asia, and the United States. To our knowledge, ForceSense remains the only physical trainer on the market that uses both Time, Motion, and Force parameters for objective assessment of laparoscopic skills.

Limitations and strengths

Few limitations were encountered while conducting this study. First, since training was allowed to continue after passing the training goals, no restrictions on the number of repetitions were imposed within the 3-week time frame. This resulted in different exposure to the offered task set and should be taken into account when comparing pre- and post-course outcomes. Secondly, we included first-year surgical residents who continued their clinical work and, although limited, still performed laparoscopic surgery during this curriculum. This might have influenced the LC and post-course outcomes. Additionally, one might say these outcomes of fundamental skills trained on elementary tasks can yet be marginally transferred to laparoscopic surgery in the OR. However, we like to emphasize that the different fundamental skills as described earlier[5–8,54] are trained and assessed sufficiently and serve as an indispensable base for further endurance and improvement in more realistic (skills center) training and eventually to perform laparoscopic surgery on real patients. Finally, it could happen that during the timespan of training, the behavior and interests of the participants change, resulting in the execution of new (surgical) activities that could also influence skills acquisition or surgical performance. To gain insight into possible relationships between activities and skills development, daily activities of participants should be monitored. Further studies could also include other physical parameters like body mass index (BMI), posture, or even hand size to gain insight into deviating data.

The strength of this study is the implementation of previously validated tasks, measurement systems, and protocol into clinical practice. The collection of objective data through mandatory participation for each first-year surgical resident in our training district resulted in an extensive amount of data with regards to individual outcomes, cohort behavior, task-specific characteristics, and parameter-specific characteristics. The present results prove the added value of objective force- and motion-based parameters in addition to the gold standard for assessment of skills: the objective structured assessment of technical skills.[55] Where the latter can only be performed in the OR or at a skills center staffed with trainers and senior surgeons who execute the assessment,[55] the current measurement system was able to provide feedback and assessment while the trainee autonomously overcomes the LC to optimally prepare for laparoscopic surgery. Skills were acquired without compromising duty or working hours regulations, and novice residents could optimally spend their working hours at the OR. These findings can now be incorporated into surgical skills curricula to facilitate surgical trainers to determine technical proficiency before moving on to a more complex simulation or the OR. As this study concerns the fundamental part of the laparoscopic LC, we suggest that, in addition to achieving theoretical knowledge, trainees are required to achieve a predefined level of technical proficiency in a milestone or EPA design for specific parts of the laparoscopic LC.

In conclusion, force- and motion-based assessment was successfully implemented and evaluated in a large proficiency-based curriculum. A significant variety of technical skills set and gradient of LCs among trainees has been shown. Curricular design should be optimized with an objective assessment to evaluate skills acquisition and to determine whether surgical trainees have reached the requisite level of technical skills compared to experts before operating on real patients.

Funding/Support

Not applicable.

Conflict of interest/Disclosure

Sem F. Hardon, Leonie A. van Gastel, and Freek Daams have no conflict of interest or financial ties to declare. Tim Horeman is the founder of MediShield B.V., a TU Delft University of Technology MedTech startup, but did not receive any financial support for this research.

Acknowledgments

The authors would like to sincerely thank Mr. R.P.M. de Hoon (Dept. of Surgery, Amsterdam UMC, Amsterdam, The Netherlands) for his skills training philosophy and his unwavering commitment to the enhancement of laparoscopic skills training in the Netherlands.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [https://doi.org/10.1016/j.surg.2021.04.024].

References

1. Aggarwal R, Darzi A. From scalpel to simulator: a surgical journey. *Surgery*. 2009;145:1–4.
2. Jakimowicz J, Fingerhut A. Simulation in surgery. *Br J Surg*. 2009;96:563–564.
3. Stefanidis D, Korndorffer Jr JR, Sierra R, Touchard C, Dunne JB, Scott DJ. Skill retention following proficiency-based laparoscopic simulator training. *Surgery*. 2005;138:165–170.
4. Moore MJ, Bennett CL. The learning curve for laparoscopic cholecystectomy. The southern surgeons club. *Am J Surg*. 1995;170:55–59.
5. Aggarwal R, Moorthy K, Darzi A. Laparoscopic skills training and assessment. *Br J Surg*. 2004;91:1549–1558.
6. Fried GM, Gill H. Surgery through the keyhole: a new view of an old art. *Mcgill J Med*. 2007;10:140–143.
7. Khan AF, Macdonald MK, Streutker C, Rowsell C, Drake J, Grantcharov T. Defining the relationship between compressive stress and tissue trauma during laparoscopic surgery using human large intestine. *IEEE J Transl Eng Health Med*. 2019;7, 3300108.
8. van der Meijden OA, Schijven MP. The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: a current review. *Surg Endosc*. 2009;23:1180–1190.
9. Stefanidis D, Sevdalis N, Paige J, et al. Simulation in surgery: what's needed next? *Ann Surg*. 2015;261:846–853.
10. Zendejas B, Brydges R, Hamstra SJ, Cook DA. State of the evidence on simulation-based training for laparoscopic surgery: a systematic review. *Ann Surg*. 2013;257:586–593.
11. Rosen J, Solazzo M, Hannaford B, Sinanan M. Objective laparoscopic skills assessments of surgical residents using Hidden Markov Models based on haptic information and tool/tissue interactions. *Stud Health Technol Inform*. 2001;81:417–423.
12. Moorthy K, Munz Y, Dosis A, Bello F, Chang A, Darzi A. Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity. *Surg Endosc*. 2004;18:1608–1612.
13. Dauster B, Steinberg AP, Vassiliou MC, et al. Validity of the MISTELS simulator for laparoscopy training in urology. *J Endourol*. 2005;19:541–545.
14. Hiemstra E, Chmarra MK, Dankelman J, Jansen FW. Intracorporeal suturing: economy of instrument movements using a box trainer model. *J Minim Invasive Gynecol*. 2011;18:494–499.
15. Van Sickle KR, McClusky 3rd DA, Gallagher AG, Smith CD. Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc*. 2005;19:1227–1231.
16. Horeman T, Dankelman J, Jansen FW, van den Dobbelsteen JJ. Assessment of laparoscopic skills based on force and motion parameters. *IEEE Trans Biomed Eng*. 2014;61:805–813.

17. Hardon SF, van Kasteren RJ, Dankelman J, Bonjer HJ, Tuynman JB, Horeman T. The value of force and torque measurements in transanal total mesorectal excision (TaTME). *Tech Coloproctol*. 2019;23:843–852.
18. Hardon SF, Horeman T, Bonjer HJ, Meijerink W. Force-based learning curve tracking in fundamental laparoscopic skills training. *Surg Endosc*. 2018;32:3609–3621.
19. Louridas M, Szasz P, de Montbrun S, Harris KA, Grantcharov TP. Can we predict technical aptitude? A systematic review. *Ann Surg*. 2016;263:673–691.
20. Louridas M, Szasz P, Fecso AB, et al. Practice does not always make perfect: need for selection curricula in modern surgical training. *Surg Endosc*. 2017;31:3718–3727.
21. van Empel PJ, van Rijssen LB, Commandeur JP, et al. Validation of a new box trainer-related tracking device: the TrEndo. *Surg Endosc*. 2012;26:2346–2352.
22. Schreuder HW, van den Berg CB, Hazebroek EJ, Verheijen RH, Schijven MP. Laparoscopic skills training using inexpensive box trainers: which exercises to choose when constructing a validated training course. *BJOG*. 2011;118:1576–1584.
23. Peters JH, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery*. 2004;135:21–27.
24. Horeman T, van Delft F, Blikkendaal MD, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Learning from visual force feedback in box trainers: tissue manipulation in laparoscopic surgery. *Surg Endosc*. 2014;28:1961–1970.
25. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg*. 1998;175:482–487.
26. Horeman T, Rodrigues SP, Jansen FW, Dankelman J, van den Dobbelsteen JJ. Force Parameters for Skills Assessment in Laparoscopy. *IEEE Trans Haptics*. 2012;5:312–322.
27. Hashimoto DA, Sirimanna P, Gomez ED, et al. Deliberate practice enhances quality of laparoscopic surgical performance in a randomized controlled trial: from arrested development to expert performance. *Surg Endosc*. 2015;29:3154–3162.
28. Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. Statistical assessment of the learning curves of health technologies. *Health Technol Assess*. 2001;5:1–79.
29. Rodrigues SP, Horeman T, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Tying different knots: what forces do we use? *Surg Endosc*. 2015;29:1982–1989.
30. Susmitha WK, Mathew G, Devasahayam SR, Perakath B, Velusamy SK. Factors influencing forces during laparoscopic pinching: Towards the design of virtual simulator. *Int J Surg*. 2015;18:211–215.
31. Hopper AN, Jamison MH, Lewis WG. Learning curves in surgical practice. *Postgrad Med J*. 2007;83:777–779.
32. Szasz P, Louridas M, Harris KA, Aggarwal R, Grantcharov TP. Assessing technical competence in surgical trainees: a systematic review. *Ann Surg*. 2015;261:1046–1055.
33. Rodrigues SP, Horeman T, Sam P, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Influence of visual force feedback on tissue handling in minimally invasive surgery. *Br J Surg*. 2014;101:1766–1773.
34. Horeman T, Meijer EJ, Harlaar JJ, Lange JF, van den Dobbelsteen JJ, Dankelman J. Force sensing in surgical sutures. *PLoS One*. 2013;8, e84466.
35. van Empel PJ, Commandeur JP, van Rijssen LB, et al. Learning curve on the TrEndo laparoscopic simulator compared to an expert level. *Surg Endosc*. 2013;27:2934–2939.
36. Horeman T, Rodrigues SP, Jansen FW, Dankelman J, van den Dobbelsteen JJ. Force measurement platform for training and assessment of laparoscopic skills. *Surg Endosc*. 2010;24:3102–3108.
37. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg*. 2004;240:518–525:discussion 25–28.
38. Vassiliou MC, Ghitulescu GA, Feldman LS, et al. The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. *Surg Endosc*. 2006;20:744–747.
39. Scott DJ, Ritter EM, Tesfay ST, Pimentel EA, Nagji A, Fried GM. Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. *Surg Endosc*. 2008;22:1887–1893.
40. Fried GM. FLS assessment of competency using simulated laparoscopic tasks. *J Gastrointest Surg*. 2008;12:210–212.
41. El-Beheiry M, McCreery G, Schlachta CM. A serious game skills competition increases voluntary usage and proficiency of a virtual reality laparoscopic simulator during first-year surgical residents' simulation curriculum. *Surg Endosc*. 2017;31:1643–1650.
42. Gallagher AG, McClure N, McGuigan J, Crothers I, Browning J. Virtual reality training in laparoscopic surgery: a preliminary assessment of minimally invasive surgical trainer virtual reality (MIST VR). *Endoscopy*. 1999;31:310–313.
43. Hagelsteen K, Sevonius D, Bergenfelz A, Ekelund M. Simball box for laparoscopic training with advanced 4D motion analysis of skills. *Surg Innov*. 2016;23:309–316.
44. Lehmann KS, Holmer C, Gillen S, et al. Suitability of a virtual reality simulator for laparoscopic skills assessment in a surgical training course. *Int J Colorectal Dis*. 2013;28:563–571.
45. Nugent E, Shirilla N, Hafeez A, et al. Development and evaluation of a simulator-based laparoscopic training program for surgical novices. *Surg Endosc*. 2013;27:214–221.
46. Luursema JM, Rovers MM, Groenier M, van Goor H. Performance variables and professional experience in simulated laparoscopy: a two-group learning curve study. *J Surg Educ*. 2014;71:568–573.
47. Kowalewski TM, White LW, Lendvay TS, et al. Beyond task time: automated measurement augments fundamentals of laparoscopic skills methodology. *J Surg Res*. 2014;192:329–338.
48. Buckley CE, Kavanagh DO, Nugent E, Ryan D, Traynor OJ, Neary PC. Zone calculation as a tool for assessing performance outcome in laparoscopic suturing. *Surg Endosc*. 2015;29:1553–1559.
49. Allen B, Nistor V, Dutson E, Carman G, Lewis C, Faloutsos P. Support vector machines improve the accuracy of evaluation for the performance of laparoscopic training tasks. *Surg Endosc*. 2010;24:170–178.
50. Cundy TP, Thangaraj E, Rafii-Tari H, et al. Force-Sensing enhanced simulation environment (ForSense) for laparoscopic surgery training and assessment. *Surgery*. 2015;157:723–731.
51. Heijnsdijk EA, Pasdeloup A, van der Pijl AJ, Dankelman J, Gouma DJ. The influence of force feedback and visual feedback in grasping tissue laparoscopically. *Surg Endosc*. 2004;18:980–985.
52. Tholey G, Desai JP, Castellanos AE. Force feedback plays a significant role in minimally invasive surgery: results and analysis. *Ann Surg*. 2005;241:102–109.
53. Horeman T. *Force-based assessment of tissue handling skills (Doctoral dissertation)*. Delft University of Technology; 2014.
54. Bonrath EM, Weber BK, Fritz M, et al. Laparoscopic simulation training: Testing for skill acquisition and retention. *Surgery*. 2012;152:12–20.
55. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84:273–278.