Development of a proficiency-based virtual reality simulation training curriculum for laparoscopic appendicectomy

Journal:	*ANZ Journal of Surgery*
Manuscript ID	ANS-2017-00176
Manuscript Type:	Original Article
Date Submitted by the Author:	20-Feb-2017
Complete List of Authors:	Sirimanna, Pramudith; University of Sydney Sydney Medical School, Academic Colorectal Unit Gladman, Marc; Concord Repatriation General Hospital, Colorectal Surgery
General Key Words:	General Surgery
Specialty Key Words:	Proficiency training, Surgical education, Virtual reality simulation, Technical skills curriculum

Title: **Development of a proficiency-based virtual reality simulation training curriculum for laparoscopic appendicectomy**

Running title: Proficiency training for appendicectomy

Authors Pramudith Sirimanna MBBS[1] Marc A Gladman PhD MRCOG FRCS (Gen Surg) FRACS [1]

Institution: [1]Academic Colorectal Unit, Sydney Medical School - Concord, University of Sydney, Australia

Correspondence: Professor Marc Gladman
1st Floor Clinical Sciences Building
Building 20,
Concord Repatriation General Hospital
Hospital Road, Concord
NSW 2139
Australia
Tel: +61 2 9767 6928
Fax: +61 2 9767 7115
E-mail: m.a.gladman@sydney.edu.au

Disclosure, Funding and Conflicts of interest:

Pramudith Sirimanna is funded by a National Health and Medical Research Council (NHMRC) postgraduate scholarship

(Scholarship number GNT1093784)

Marc Gladman has received support from Karl Storz and Applied medical, during the conduct of the study and personal fees from Health Ed, Elsevier, and Abbvie LTD, outside the submitted work

SLHD Human Research Ethics Committee approval code: AU/6/DC6519

This paper was previously presented as an oral presentation at the 10th Annual Academic Surgical Congress (ASC), February 3-5 2015 in Las Vegas, NV.

Word count 3887

Body of manuscript including abstract / references	2321
5x tables / figures at 250 words each	1250

Additional results and tables x2 and figures x3 included as supplementary files for online only

ABSTRACT

Introduction: Proficiency-based virtual reality (VR) training curricula improve intra-operative performance but there is a paucity of curricula for LA. This study aimed to develop an evidence-based training curriculum for laparoscopic appendicectomy (LA).

Methods: 10 experienced (>50 LAs), 8 intermediate (10-30 LAs) and 20 inexperienced (<10 LAs) operators performed guided and unguided LA tasks on a high-fidelity VR simulator using internationally relevant techniques. Ability to differentiate levels of experience (construct validity) was measured using simulator-derived metrics. Learning curves were analyzed. Proficiency benchmarks were defined by performance of the experienced group. Intermediate and experienced participants completed a questionnaire to evaluate the realism (face validity) and relevance (content validity).

Results: Sixteen of 18 (89%) surgeons considered the VR model to be visually realistic and 17 (95%) that it was representative of actual practice. All 'guided' modules demonstrated construct validity (P<0.05), with learning curves that plateaued between sessions 6 and 9 (P<0.01). When comparing inexperienced to intermediates to experienced, the 'unguided' LA module demonstrated construct validity for economy of motion (5.00 vs 7.17 vs 7.84, respectively, P<0.01) and task time (864.5s vs 477.2s vs 352.1s, respectively, P<0.01). Construct validity was also confirmed for number of movements, path length and idle time. Validated modules were used for curriculum construction with proficiency benchmarks used as performance goals.

Discussion: A VR LA model was realistic and representative of actual practice and was validated as a training and assessment tool. Consequently, the first evidence-based internationally applicable training curriculum for LA was constructed that facilitates skill acquisition to proficiency.

KEY WORDS

Laparoscopic appendicectomy

Proficiency training

Surgical education

Technical skills curriculum

Virtual reality simulation

INTRODUCTION

Surgical training is associated with a steep learning curve, which may be associated with errors, complications and mortality [1, 2]. Adverse events may occur in up to 10% of all hospital admissions with two-thirds occuring within the surgical domain [3 4 5 6 7]; half of which are preventable and attributable to technical errors [8 5 9]. Recently, there has been interest in virtual reality (VR) simulation, as this allows technical skill acquisition with improved actual OT performance [10 11] and creation of structured training curricula using expert benchmarks of skill i.e. competency-based performance goals [12 13, 14]. Such curricula can improve trainee performance in the actual OT [15-17].

Appendicectomy is the most common emergency operation and is often performed laparoscopically due to documented benefits [18 19]. However, laparoscopic appendicectomy (LA) requires specialized skills resulting in a notable learning curve of up to 30 cases [20]. Despite this, LA remains the principle index operation for trainees; often being the first experience of laparoscopic surgery. The aims of this study were to demonstrate that a VR simulation model of LA is (i) realistic (face valid), (ii) relevant (content valid) to clinical practice and (iii) useful as a training and assessment tool for LA by demonstrating its ability to improve novice surgeons' performance and to differentiate between levels of experience (construct validity). Finally, a structured, proficiency-based VR training curriculum was developed.

METHODS

Participant selection

Subjects were stratified according to their degree of experience of LA as follows: experienced (performed >50 LAs), intermediate (10-30 LAs) and inexperienced (<10 LAs) operators. Inexperienced subjects who had not performed were asked to watch three videos of LAs. As previously demonstrated, a minimum of eight participants per group was required [13]. Individuals with previous laparoscopic simulation training were excluded from the study. Ethical approval was granted by the Human Research Ethics Committee –Sydney Local Health District - Concord (Approval code: AU/6/DC6519). All subjects provided informed consent to participate.

Virtual reality simulation tool

The LA training tool of the LAP Mentor™ VR laparoscopic surgical simulator (Simbionix Corporation, Cleveland, Ohio, USA) was used. In addition to nine previously validated basic laparoscopic skills tasks [13], the LA tool consists of five 'guided' procedural tasks and an 'unguided' full LA task (supplementary Figure 1). For this study the 'unguided' full LA task with a mildly inflamed appendix in a pelvic position was used. A detailed description of the simulated LA tasks is specified in supplementary table 1.

Tasks performed

All participants underwent baseline skills testing [13]. Subsequently, both intermediate and experienced operators performed each of the five 'guided' procedural tasks on 2 occasions (Figure 1). Each intermediate and experienced operator completed the 'unguided' full LA task on 2 occasions (Figure 1). To reflect international variation, LA was performed using three different methods. Technique 1: clips to control the appendicular artery and endoloops to divide the appendix; technique 2: an energy device (Harmonic scalpel™, Ethicon, Cincinnati, OH) to control the appendicular artery and a linear stapler to divide the appendix; and technique 3: a stapler to control the appendicular artery and divide the appendix (Figure 1). Intermediate and experienced participants completed a 23-item questionnaire (supplementary Table 2) to evaluate the face and content validity of the VR simulation model, using a 5-point Likert scale.

Inexperienced operators were randomized into two groups (group A and group B) using a sealed envelope technique. Group A conducted ten repetitions of the five 'guided' LA procedural tasks, whereas group B conducted ten repetitions of the full 'unguided' LA task using technique 1 (Figure 1). Additionally, during the first and tenth session, group B conducted two further full 'unguided' LAs using technique 2 and 3.

Data and statistical analyses

The VR simulator objectively measures total: task time (TT), number of movements (NOM), economy of motion (EOM), path length (PL) and idle time (IT). The median performance during the second session was compared using Kruskal-Wallis or Mann-Whitney *U* tests to determine whether the model could differentiate between varying levels of experience (construct validity) and authenticate its use as an *assessment* tool. Non-parametric repeated measures analysis of variance (Friedman) test assessed learning curves in the inexperienced group to determine whether performance improved with repeated practice to substantiate the simulator as a *training* tool. For each task, those with the longest learning curve (i.e. greatest numbers of sessions taken for a plateau to be obtained) were deemed to be the most 'challenging' modules. Benchmarks of proficiency were defined as the median scores obtained for each simulated task by experienced surgeons. Finally, the proportion of intermediate and experienced surgeon responders who agreed or strongly agreed with each item on the post-study questionnaire was calculated to determine the face and content validity of the VR simulation model. Data were analyzed with SPSS® version 22 (SPSS, Chicago, Illinois, USA) using non-parametric tests. A P value of < 0.05 was considered statistically significant.

Curriculum construction

A proficiency-based training curriculum for LA was constructed using each simulated LA module that demonstrated construct validity and learning curves. The most challenging tasks were used for summative assessment at each step of the curriculum. Proficiency criteria, defined by benchmarks of experienced surgeons' performance, were used as performance goals during these summative assessments.

RESULTS

Ten experienced, eight intermediate and 20 inexperienced operators completed the study.

Face and content validity

See supplementary results (Doc S1).

Construct validity: 'Guided' Tasks

All five 'guided' tasks demonstrated construct validity as evidenced by significant performance differences between the three groups for NOM, PL, IT and TT (P<0.05) (see supplementary Figure 2) (Table 1). However, differences were only observed between the three levels of experience for EOM during 'guided' tasks 3 and 5 (P<0.05) but not for 1, 2 and 4 (Table 1).

When comparing the inexperienced and experienced groups, performance differences were demonstrated during all 'guided' tasks for NOM, PL, IT and TT (P<0.05). Furthermore, differences in EOM were observed between these groups during 'guided' tasks 3 and 5 (P<0.05). Similar performance differences were observed between the inexperienced and intermediate groups for NOM, PL, IT and TT during 'guided' task 1 and 4 (P<0.05), for IT during 'guided' task 2 (P<0.05), for EOM and ID during 'guided' task 3 (P<0.05) and for IT and TT during 'guided' task 5 (P<0.05). Lastly, significant differences were observed between the intermediate and experienced groups for NOM, PL, IT and TT during 'guided' tasks 2, 3 and 5 (P<0.05) and for NOM, IT and TT during 'guided' task 1 (P<0.05). No differences in performance were observed between the intermediate and experienced groups during 'guided' task 4.

Construct validity: 'Unguided' Tasks

All three 'unguided' tasks demonstrated construct validity with significant performance differences between the three groups for EOM, NOM, PL, IT and TT (P<0.01) (Table 1). Significant differences in performance were observed during the three 'unguided' task attempts for all metrics when comparing the inexperienced group to the experienced and intermediate groups (P<0.05). Intermediate group performance differed from that of the experienced group for NOM, PL, IT and TT during all three 'unguided' task attempts (P<0.05) but not for EOM.

Learning curves: 'Guided' tasks

Significant learning curves were demonstrated for the inexperienced group for EOM, NOM, PL, IT and TT during all 'guided' tasks (P<0.01) (Figure 2). A plateau in performance was reached during the 6th session for 'guided' task 4, during the 8th session for 'guided' tasks 1, 3 and 5 and during the 9th session for 'guided' task 2.

Learning curves: 'Unguided' tasks

Significant learning curves were demonstrated for the inexperienced group for EOM, NOM, PL, IT and TT during 'unguided' task using technique 1 (P<0.01). However, no plateau in performance was reached for TT. Statistically significant differences in inexperienced group performance were demonstrated between the 1st and 10th sessions for all five simulator-derived metrics during 'unguided' task attempts using technique 2 and 3.

Proficiency criteria and Curriculum construction

The proficiency benchmarks for each of the 'guided' and 'unguided' tasks and a summary of the results for the tasks to be used for summative assessment during curriculum construction are summarized in table 2. These, in addition to the other validated tasks, were used to develop a proficiency-based VR technical skills curriculum for LA (Figure 3).

DISCUSSION

This study validated a VR simulation model of LA as a training and assessment tool. Face and content validity were demonstrated and all VR simulation tasks were shown to be construct valid with demonstrable differences in performance between the three levels of experience. Specifically, there were significant differences between the three groups for all five of the guided tasks and all three of the unguided tasks. Additionally, learning curves were demonstrable for each task to illustrate that repetitive practice improved the performance of inexperience surgeons with guided task 4 ("division of the mesoappendix and base of appendix using a stapler") proving the easiest and task 2 ("dissecting the mesoappendix and clipping the artery") the hardest. Accordingly, a proficiency-based curriculum was constructed using these findings and benchmarks of proficient performance obtained from experienced surgeons.

The curriculum enables novice surgeons to practice the skills required to perform a LA in a stepwise manner, with advancement through the steps only once proficiency is attained. Two repetitions of the five 'guided' LA tasks are performed followed by training using the two most challenging 'guided' LA tasks with the longest learning curves, i.e. task 3 ("clipping the artery and ligating the appendix using a loop") and 5 ("control of the artery using energy and ligation of appendix using loops"). 'Guided' task 1 ("dissection of mesenteric window") and 2 ("dissecting the mesoappendix and clipping the artery") were not used as these 'steps' are contained within tasks 3 and 5. Following attainment of proficiency in these 'guided' tasks, two repetitions of the 'unguided' full LA task are performed using each of the three techniques. Completion of the curriculum occurs when proficiency is achieved for the 'unguided' full LA task using technique 1 (clips/endoloops) and 3 (stapler) (Figure 3). Unguided technique 1 was included as it was judged to be most difficult given that no maximum plateau in learning occurred and the 'unguided' task using technique 3 was used as intermediate surgeons performed worse than experienced surgeons. Attainment of proficiency at each stage of the curriculum must be demonstrated at two consecutive sessions in order to negate the possibility of achieving the proficiency scores by chance. Finally, it is recommended that a maximum of two sessions be allowed per day (each at least one hour apart), to ensure adherence to the principle of 'distributed' learning [21 22].

A number of studies have developed proficiency-based training curricula for surgery [12 13 14]. Indeed, a novel training pathway for the management of appendicitis has recently been developed, which included a proficiency-based VR training curriculum for LA adapted from a previously developed curriculum [27 28]. Despite the potential benefits of VR training, its uptake into surgical training has been poor, possibly due to concerns over expense and/or difficulties incorporating into the schedules of trainees. However, it has recently been demonstrated that VR simulation training is more cost effective than conventional surgical training and box training for programs with more than 10 residents[28]. The tasks in the presented LA curriculum take approximately 12 to 20 minutes and can be performed out of the OT, which may be of particular relevance in the future given the recent mandatory restrictions to maximum working hours in North America, Europe and Queensland.

Limitations of this study include potential selection bias from the recruitment strategy that may have favored surgically inclined, well-motivated novice trainees. Further, all experienced surgeons were recruited from a single-centre teaching hospital. The comparison of performance of the novice to experienced groups may introduce bias, as the range of abilities that occur within each group may not be well represented. However, the recruitment of intermediate surgeons was from a larger network of surgical trainees, and the performance differences observed between groups provides evidence of the capability of the simulation tool to discriminate despite this. Alternative systems for validation exist, including Messick's validity framework [23 24], but the framework used within this study has previously been utilized in the production of curricula that improves actual operative performance [14, 15 17, 25], including the Foundations of Laparoscopic Surgery (FLS) training curriculum endorsed by the American College of Surgeons.[26]

CONCLUSION

This study describes an internationally applicable proficiency-based virtual reality technical skills curriculum for LA. Although this curriculum is not designed to substitute skills acquisition in the operating theatre, it provides a useful adjunct to obtain key skills required for LA in a risk-free environment. Ultimately, it is hoped that the curriculum will improve intra-operative surgical performance via the creation of a 'pre-trained novice'.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contributions of A/ Professor Rajesh Aggarwal and A/ Professor Vincent Lam for advice regarding study design and participant recruitment, respectively.

REFERENCES

[1] Reichenbach DJ, Tackett AD, Harris J, et al. Laparoscopic colon resection early in the learning curve: what is the appropriate setting? *Ann Surg*. 2006; 243:730-5; discussion 5-7.

[2] Watson DI, Baigrie RJ, Jamieson GG. A learning curve for laparoscopic fundoplication. Definable, avoidable, or a waste of time? *Ann Surg*. 1996; 224:198-203.

[3] Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *The New England journal of medicine*. 1991; 324:370-6.

[4] Wilson RM, Harrison BT, Gibberd RW, Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Med J Aust*. 1999; 170:411-5.

[5] Gawande AA, Thomas EJ, Zinner MJ, Brennan TA. The incidence and nature of surgical adverse events in Colorado and Utah in 1992. *Surgery*. 1999; 126:66-75.

[6] Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review.[Erratum appears in BMJ 2001 Jun 9;322(7299):1395]. *Bmj*. 2001; 322:517-9.

[7] Baker GR, Norton PG, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *Cmaj*. 2004; 170:1678-86.

[8] Thomas EJ, Studdert DM, Newhouse JP, et al. Costs of medical injuries in Utah and Colorado. *Inquiry*. 1999; 36:255-64.

[9] Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *New England Journal of Medicine*. 1991; 324:377-84.

[10] Gurusamy KS, Aggarwal R, Palanivelu L, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. *Cochrane Database Syst Rev*. 2009:CD006575.

[11] 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; 257:586-93.

[12] Aggarwal R, Grantcharov TP, Eriksen JR, et al. An evidence-based virtual reality training program for novice laparoscopic surgeons. *Ann Surg*. 2006; 244:310-4.

[13] Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A. Development of a virtual reality training curriculum for laparoscopic cholecystectomy. *Br J Surg*. 2009; 96:1086-93.

[14] Palter VN, Graafland M, Schijven MP, Grantcharov TP. Designing a proficiency-based, content validated virtual reality curriculum for laparoscopic colorectal surgery: a Delphi approach. *Surgery*. 2012; 151:391-7.

[15] Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg*. 2007; 193:797-804.

[16] Palter VN, Grantcharov TP. Development and validation of a comprehensive curriculum to teach an advanced minimally invasive procedure: a randomized controlled trial. *Ann Surg*. 2012; 256:25-32.

[17] Palter VN, Orzech N, Reznick RK, Grantcharov TP. Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery: a randomized controlled trial. *Ann Surg*. 2013; 257:224-30.

[18] Saia M, Buja A, Baldovin T, et al. Trend, variability, and outcome of open vs. laparoscopic appendectomy based on a large administrative database. *Surgical endoscopy*. 2012; 26:2353-9.

[19] National Surgical Research C. Multicentre observational study of performance variation in provision and outcome of emergency appendicectomy. *Br J Surg*. 2013; 100:1240-52.

[20] Kim SY, Hong SG, Roh HR, Park SB, Kim YH, Chae GB. Learning curve for a laparoscopic appendectomy by a surgical trainee. *Journal of the Korean Society of Coloproctology*. 2010; 26:324-8.

[21] Mackay S, Morgan P, Datta V, Chang A, Darzi A. Practice distribution in procedural skills training: a randomized controlled trial. *Surgical endoscopy*. 2002; 16:957-61.

[22] Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. *Ann Surg*. 2006; 244:400-9.

[23] Messick S. *Validity*. 3rd ed: New York: American Council on Education and Macmillan., 1989.

[24] Cook DA, Zendejas B, Hamstra SJ, Hatala R, Brydges R. What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. *Adv Health Sci Educ Theory Pract*. 2014; 19:233-50. doi: 10.1007/s10459-013-9458-4. Epub 2013 May 2.

[25] Aggarwal R, Ward J, Balasundaram I, Sains P, Athanasiou T, Darzi A. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg*. 2007; 246:771-9.

[26] 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; 135:21-7.

[27] Beyer-Berjot L, Patel V, Acharya A, et al. Surgical training: Design of a virtual care pathway approach. *Surgery*. 2014; 156:689-97.

[28] Orzech N, Palter VN, Reznick RK, Aggarwal R, Grantcharov TP. A comparison of 2 ex vivo training curricula for advanced laparoscopic skills: a randomized controlled trial. *Ann Surg*. 2012; 255:833-9.

FIGURE LEGENDS

Figure 1: **Protocol of study** Intermediate and experienced surgeons completed the five 'Guided' tasks followed by the 'Unguided' tasks. Inexperienced surgeons were stratified into group A, who conducted 10 repetitions of the 'Guided' tasks, and group B who conducted 10 repetitions of an 'Unguided' task.

Figure 2: **Learning curve for total task time for 'guided' task 1.** Horizontal lines within boxes, and whiskers represent median, interquartile range and range respectively. It can be seen that the learning curve plateaus at the 8th session.

Figure 3: **Evidence-based VR curriculum to train technical skills required for laparoscopic appendicectomy to proficiency.** The curriculum presented enables novice surgeons to practice the skills required for LA in a stepwise manner, with advancement only allowed to occur once proficiency is attained.

List of Supporting Information

Supplementary Doc 1 **Supplementary Result**

Supplementary Figure 1: **Screen shots obtained from the LAP Mentor™ VR laparoscopic surgical simulator (Simbionix Corporation, Cleveland, Ohio, USA)** (a) 'Guided' task 1: Dissecting the mesenteric window (b) 'Unguided' full laparoscopic appendicectomy task

Supplementary Figure 2: **Total task time for 'guided' task 1:** Horizontal lines within boxes, and whiskers represent median, interquartile range and range respectively. The total task time taken was significantly different between the 3 groups with the inexperienced taking longest and the experienced being the fastest, and the intermediate group's performance fell in between.

Table 1: Construct valid metrics for 'guided' and 'unguided' tasks

		Simulator-derived Metrics				
Task	**Group**	**Total economy of motion**	**Total number of movements**	**Total path length**	**Idle time**	**Total task time**
Guided Task 1	Inexperienced	4.75	**651***	**790.03cm***	**124.49s***	**412.67s***
	Intermediate	4.92	**234.5***	**379.04cm***	**36.96s***	**143.34s***
	Experienced	5.83	**149.5***	**261.63cm***	**20.92s***	**78.38s***
Guided Task 2	Inexperienced	5.01	**408***	**708.06cm***	**99.83s***	**314.78s***
	Intermediate	5.34	**398***	**675.46cm***	**71.74s***	**248.84s***
	Experienced	5.85	**264***	**430.2cm***	**51.97s***	**177.21s***
Guided Task 3	Inexperienced	**5.54†**	**852.5†**	**1593.91cm†**	**258.63s***	**645.02s***
	Intermediate	**6.33†**	**891†**	**1797.87cm†**	**201.67s***	**593.58s***
	Experienced	**7.07†**	**670.5†**	**1235.93cm†**	**148.58s***	**427.98s***
Guided Task 4	Inexperienced	5.70	**328†**	**606.42cm†**	**56.8s***	**262.3s***
	Intermediate	6.25	**256.5†**	**491.59cm†**	**32.78s***	**179.53s***
	Experienced	6.34	**245.5†**	**474.29cm†**	**31.24s***	**175.07s***
Guided Task 5	Inexperienced	**6.23†**	**652***	**1164.26cm***	**178.76s***	**433.5s***
	Intermediate	**7.09†**	**581.5***	**1284.86cm***	**134.46s***	**388.73s***
	Experienced	**8.18†**	**373.5***	**872.85cm***	**87.72s***	**265.07s***
Unguided Task Technique 1 (Clips / Endoloops)	Inexperienced	**5.00***	**1101***	**1797.08cm***	**325.43s***	**864.49s***
	Intermediate	**7.17***	**690.5***	**1573.51cm***	**160.44s***	**477.2s***
	Experienced	**7.84***	**532***	**1315.09cm***	**118.45s***	**352.12s***
Unguided Task Technique 2 (Energy device / stapler)	Inexperienced	**5.00***	**432.5***	**648.68cm***	**88.78s***	**360.34s***
	Intermediate	**7.59***	**221.5***	**537.51cm***	**29.63s***	**146.66s***
	Experienced	**8.64***	**175***	**418.5cm***	**22.98s***	**106.61s***
Unguided Task Technique 3 (Stapler)	Inexperienced	**5.17***	**453***	**748.79cm***	**92.46s***	**384.02s***
	Intermediate	**7.34***	**245***	**536.59cm***	**28.78s***	**167.47s***
	Experienced	**7.87***	**164***	**393.04cm***	**19.59s***	**101.12s***

KEY: † Significant differences between the groups with $P < 0.05$

* Significant differences between the groups with $P < 0.01$

Table 2: Summary of results for tasks used for curriculum construction and assessment within curriculum

Task	Metric	Construct validity	Learning curve	Plateau session	Proficiency benchmark*
Guided task 3	Total economy of motion	YES	YES	8th	7.07
	Total number of movements	YES	YES	6th	670.5
	Total path length	YES	YES	6th	1235.93 cm
	Idle time	YES	YES	7th	148.58 secs
	Total task time	YES	YES	8th	427.98 secs
Guided task 5	Total economy of motion	YES	YES	8th	8.18
	Total number of movements	YES	YES	5th	373.5
	Total path length	YES	YES	5th	872.95 cm
	Idle time	YES	YES	8th	87.72 secs
	Total task time	YES	YES	8th	265.07 secs
Unguided task – Technique 1	Total economy of motion	YES	YES	6th	7.84
	Total number of movements	YES	YES	6th	532
	Total path length	YES	YES	5th	1315.09 cm
	Idle time	YES	YES	8th	118.45 secs
	Total task time	YES	YES	No plateau	352.12 secs
Unguided task – Technique 3	Total economy of motion	YES	YES	N/A	7.87
	Total number of movements	YES	YES	N/A	164
	Total path length	YES	YES	N/A	393.04 cm
	Idle time	YES	YES	N/A	19.59 secs
	Total task time	YES	YES	N/A	101.12 secs

Key: * median values from second session performed by experienced surgeons

Figure 1: Protocol of study

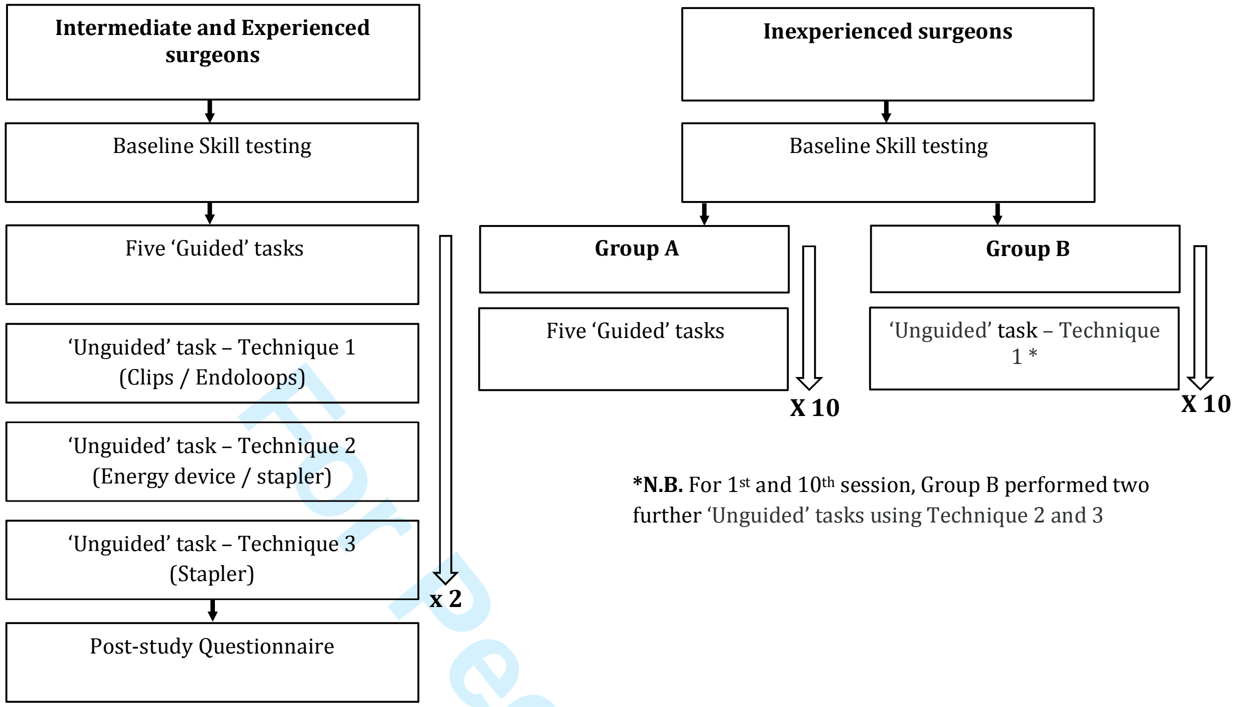

Figure 1 Legend: Above is the study protocol. Intermediate and experienced surgeons completed the five 'Guided' tasks followed by the 'Unguided' tasks. Inexperienced surgeons were stratified into group A, who conducted 10 repetitions of the 'Guided' tasks, and group B who conducted 10 repetitions of an 'Unguided' task.

Figure 2: Learning curve for total task time for 'guided' task 1

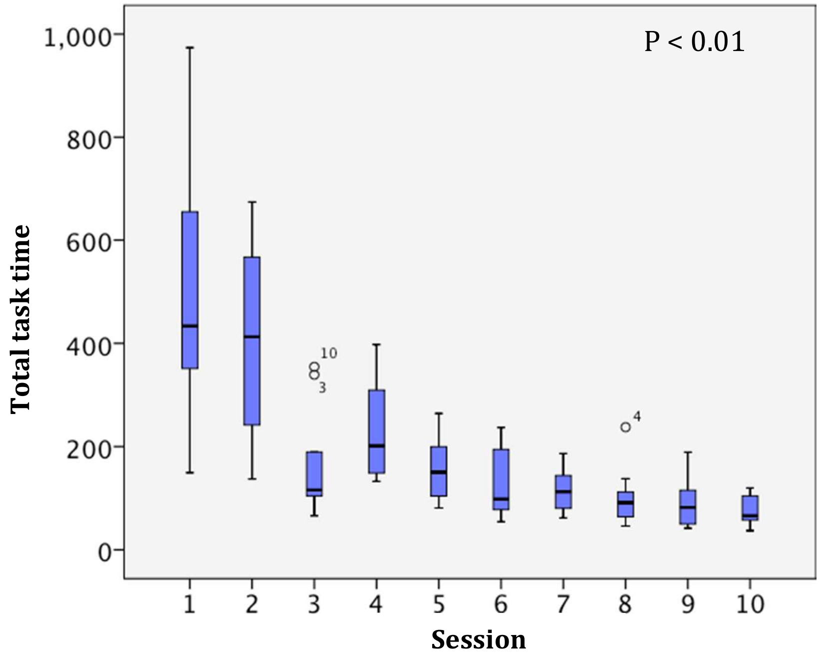

Figure 2 Legend: Horizontal lines within boxes, and whiskers represent median, interquartile range and range respectively. It can be seen that the learning curve plateaus at the 8th session.

Figure 3: Evidence-based VR curriculum to train technical skills required for laparoscopic appendicectomy to proficiency

Nine basic tasks
Nine tasks performed twice on the same day in two sessions, each session > 1 h apart

Two basic tasks (clipping and grasping, and two-hand maneuvers)
Performed for a maximum of two sessions per day, each session > 1 h apart
Completion of training when all of the following levels of skill are achieved on two consecutive sessions

Clipping and grasping
Time taken < 100 s

Two-hand maneuvers:
Total time taken < 90 s
Total no. of movements < 100
Total path length < 440 cm

Five 'Guided' Laparoscopic Appendicectomy tasks
Five 'Guided' tasks performed twice on the same day in two sessions, each session > 1 h apart

Two 'Guided' Laparoscopic Appendicectomy tasks:
Guided task 3: Clipping the Artery and Ligating the Appendix Using a Ligating Loop
Guided task 5: Control of the Artery Using Energy and Ligating the Appendix Using a Ligating Loop
Performed for a maximum of two sessions per day, each session > 1 h apart
Completion of training when all of the following levels of skill are achieved on two consecutive sessions

Guided task 3: Clipping the Artery and Ligating the Appendix Using a Ligating Loop
Total economy of motion 7.0
Total no. of movements 675
Total path length 1240 cm
Idle time 150 s
Total task time 430 s

Guided task 5: Control of the Artery Using Energy and Ligating the Appendix Using a Ligating Loop
Total economy of motion 8.0
Total no. of movements 375
Total path length 875 cm
Idle time 90 s
Total task time 270 s

Three 'Unguided' Laparoscopic Appendicectomy tasks using:
Technique 1: Clipping the Artery and Ligating the Appendix Using a Ligating Loop
Technique 2: Control of the Artery Using Energy and Division of the Appendix Using a Stapler
Technique 3: Division of the Mesoappendix and Base of the Appendix Using a Stapler
Three 'Unguided' modules performed twice on the same day in two sessions, each session > 1 h apart

Two 'Unguided' Laparoscopic Appendicectomy tasks:
Technique 1: Clipping the Artery and Ligating the Appendix Using a Ligating Loop Guided
Technique 3: Division of the Mesoappendix and Base of the Appendix Using a Stapler
Performed for a maximum of two sessions per day, each session > 1 h apart
Completion of training when all of the following levels of skill are achieved on two consecutive sessions

Unguided task - Technique 1: Clipping the Artery and Ligating the Appendix Using a Ligating Loop Guided
Total economy of motion 7.5
Total no. of movements 535
Total path length 1320 cm
Idle time 120 s
Total task time 355 s

Unguided task - Technique 3: Division of the Mesoappendix and Base of the Appendix Using a Stapler
Total economy of motion 7.5
Total no. of movements 165
Total path length 395 cm
Idle time 20 s
Total task time 105 s

Figure 3 Legend: The curriculum presented enables novice surgeons to practice the skills required for LA in a stepwise manner, with advancement only allowed to occur once proficiency is attained.

Doc 1 SUPPLEMENTARY RESULTS

Face and content validity

All 18 intermediate and experienced participants completed the post-study questionnaire. Of these, 16 (89%) agreed/strongly agreed that the VR model was visually realistic and 17 (95%) that it was representative of performing a LA. Specifically, 17 participants (95%) regarded the VR anatomy as accurate and 16 (83%) that it was visually comparable to an inflamed appendix. All participants considered the VR instruments to be visually accurate and 16 (89%) that instrument movements and camera angles were authentic. Despite only four participants (22%) reporting tactile feedback to be realistic, tissue handling and behavior was regarded as accurate by 11 (61%). Indeed, 17 participants (95%) agreed/strongly agreed that dissection and division of the mesoappendix and appendix was realistic. Finally, all participants supported the VR model as a training tool and 17 (95%) supported its use as an assessment tool.

Supplementary Table 1: Description of laparoscopic appendicectomy tasks on the LAP Mentor™ virtual reality simulator

'Guided' Tasks	Description of Task
Task 1: "Dissection of mesenteric window"	Using a Maryland dissector, dissect a window in the mesentery at the base of the appendix within the specified area
Task 2: "Dissecting the Mesoappendix and clipping the artery"	Using a Maryland dissector, dissect within a specified area of the mesoappendix to identify the appendicular artery. Clip the appendicular artery within a specified area and then safely cut between the clips.
Task 3: "Clipping the artery and ligating the appendix using a loop"	Complete tasks in Tasks 1 and 2. Then place endoloops over the appendix at specified locations and safely cut between these endoloops
Task 4: "Division of the mesoappendix and base of appendix using a stapler"	Complete task in Task 1. Using a vascular stapler, transect the mesoappendix and ligate the appendicular artery within a specified area. Then transect the appendix at the base within a specified area using a linear stapler.
Task 5: "Control of the artery using energy and ligation of appendix using loops"	Complete task in Task 1. Using an appropriate energy device, transect the mesoappendix and ligate the appendicular artery within a specified area. Then place endoloops over the appendix at specified locations and safely cut between these endoloops
'Unguided' Task	**Description of Task**
"Appendix in regular (pelvic) position"	Full virtual appendicectomy, based on anatomies created from CT/MRI real patient data. Practice a complete appendicectomy procedure, with a range of appropriate instruments.

Supplementary Table 2: Examples of items on questionnaire to assess face and content validity

Item
The virtual reality (VR) appendix was a visually realistic representation of actual appendix anatomy
Anatomical relationship between VR appendix, caecum and small bowel was realistic
The laparoscopic instruments were visually realistic in their movements
This model contained a realistic representation of the laparoscopic instruments that would be available in the actual operating theatre
Tissue handling was realistic in comparison to actual tissue
The VR tissue dissection was realistic in comparison to actual tissue
The tactile feedback was realistic in comparison to actual tissue
As a whole, the LAP Mentor VR appendicectomy model was a visually realistic model of an actual laparoscopic appendicectomy
This model as a whole was a realistic representation of performing a laparoscopic appendicectomy
This model is a useful tool for training skills required to perform a laparoscopic appendicectomy
This model is a useful tool for the assessment of trainees' performance of a laparoscopic appendicectomy

Supplementary Figure 1: Screen shots obtained from the LAP Mentor™ VR laparoscopic surgical simulator (Simbionix Corporation, Cleveland, Ohio, USA)

(a)

(b)

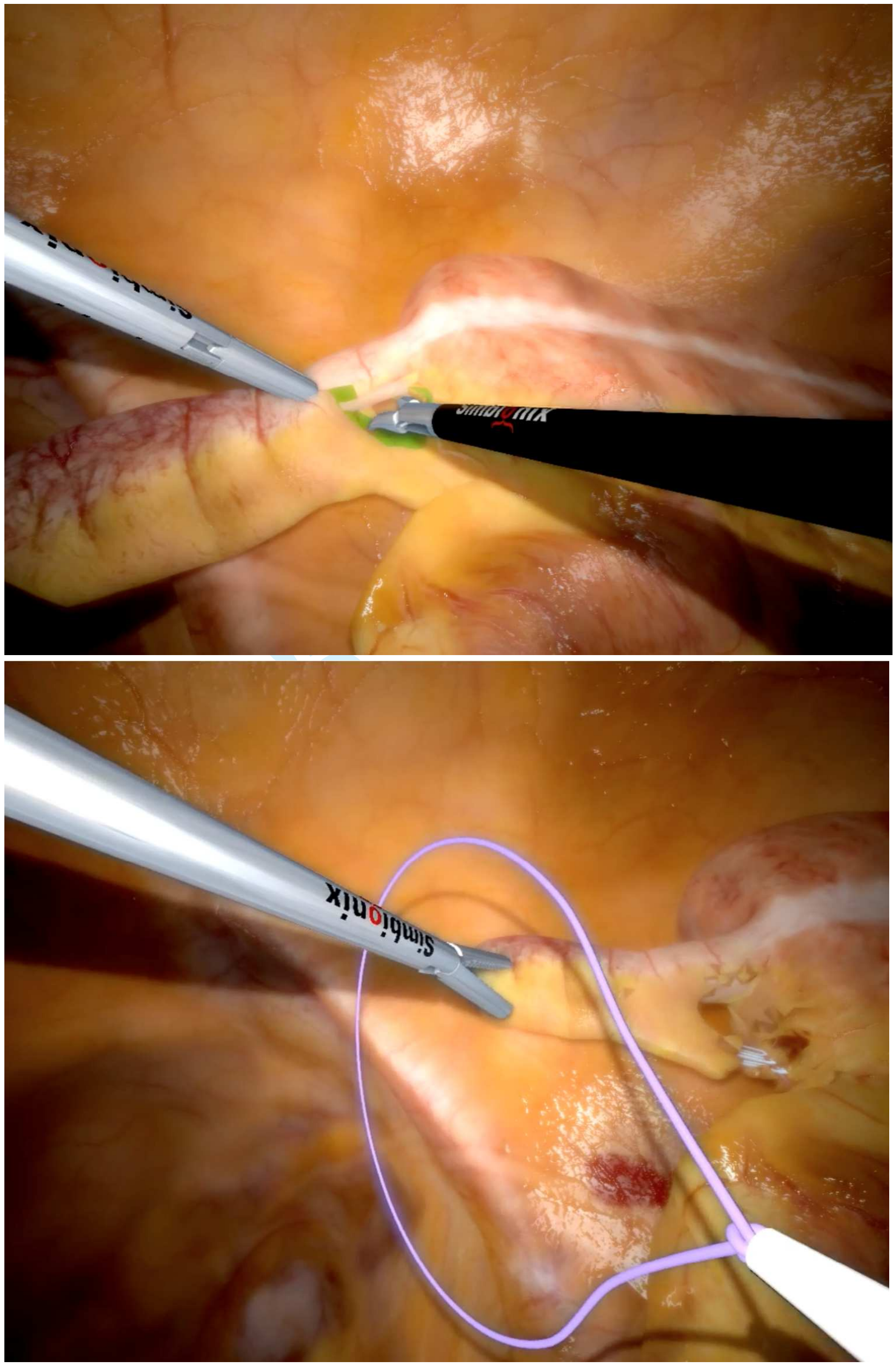

Figure Legend: (a) 'Guided' task 1: Dissecting the mesenteric window (b) 'Unguided' full laparoscopic appendicectomy task

Supplementary Figure 2: Total task time for 'guided' task 1.

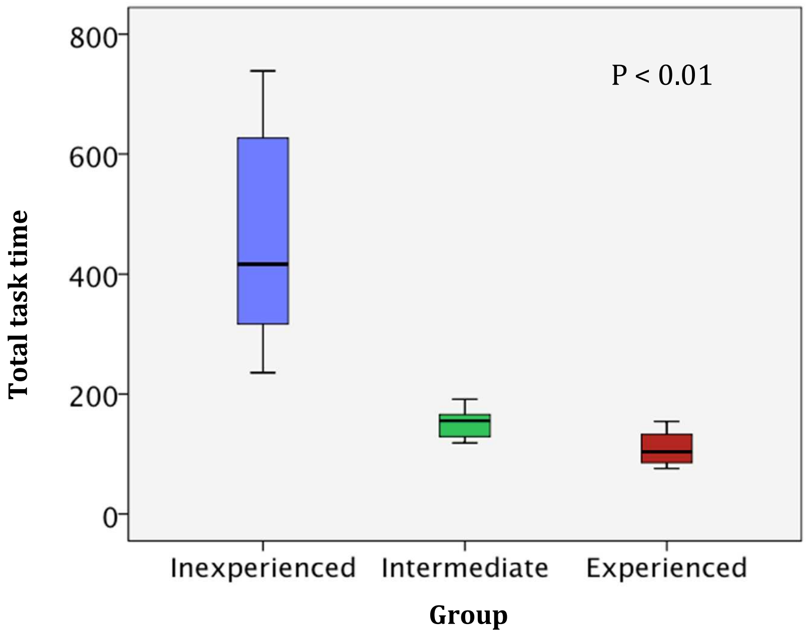

Figure Legend: Horizontal lines within boxes, and whiskers represent median, interquartile range and range respectively. The total task time taken was significantly different between the 3 groups with the inexperienced taking longest and the experienced being the fastest, and the intermediate group's performance fell in between.