Instructor Feedback Versus No Instructor Feedback on Performance in a Laparoscopic Virtual Reality Simulator

A Randomized Trial

Jeanett Strandbygaard, PhD, Flemming Bjerrum, MD,* Mathilde Maagaard, MD,* Per Winkel, DMSc,† Christian Rifbjerg Larsen, PhD,‡ Charlotte Ringsted, PhD,§ Christian Gluud, DMSc,† Teodor Grantcharov, PhD,¶ Bent Ottesen, DMSc,|| and Jette Led Sorensen, MD**

Objective: To investigate the impact of instructor feedback versus no instructor feedback when training a complex operational task on a laparoscopic virtual reality simulator.

Background: Simulators are now widely accepted as a training tool, but there is insufficient knowledge about how much feedback is necessary, which is useful for sustainable implementation.

Methods: A randomized trial complying with CONSORT Statement. All participants had to reach a predefined proficiency level for a complex operational task on a virtual reality simulator. The intervention group received standardized instructor feedback a maximum of 3 times. The control group did not receive instructor feedback. Participants were senior medical students without prior laparoscopic experience ($n = 99$). Outcome measures were time, repetitions, and performance score to reach a predefined proficiency level. Furthermore, influence of sex and perception of own surgical skills were examined.

Results: Time (in minutes) and repetitions were reduced in the intervention group (162 vs 342 minutes; $P < 0.005$) and (29 vs 65 repetitions; $P < 0.005$). The control group achieved a higher performance score than the intervention group (57% vs 49%; $P = 0.004$). Men used less time (in minutes) than women ($P = 0.037$), but no sex difference was observed for repetitions ($P = 0.20$). Participants in the intervention group had higher self-perception regarding surgical skills after the trial ($P = 0.011$).

Conclusions: Instructor feedback increases the efficiency when training a complex operational task on a virtual reality simulator; time and repetitions used to achieve a predefined proficiency level were significantly reduced in the group that received instructor feedback compared with the control group. Trial registration number: NCT01497782.

Keywords: feedback, laparoscopy, predefined proficiency level, training, virtual reality simulation

(*Ann Surg* 2013;257: 839–844)

For virtual reality simulation, the benefits are clear; the drawbacks are less clear. Throughout the last decade, several studies have found a positive effect of surgical virtual reality training on the learning curve and improvement of basic psychomotor skills in the operating room.[1–5] Despite the now well-established advantages of virtual reality simulators, most surgical and gynecological departments encounter hurdles when implementing simulator training in surgical practice.[6] This is mainly due to concern about the time and human resources needed to train novice surgeons to an adequate level,[7] along with a lack of knowledge on how to design a training program.[8] Questions have especially been asked regarding frequency, amount, and type of feedback (ie, simple vs escalating feedback) to obtain the best learning outcomes in complex operational virtual reality tasks.[9–12]

Feedback can be defined as the provision or return of performance-related information to the performer and is an important part of learning in medical education.[13] In the laparoscopic virtual reality setting, no studies or trials have investigated how instructor feedback affects learning to a predefined proficiency level in complex operational tasks; however, in more basis tasks, such as coordination and instrument navigation, no advantages of instructor feedback have been found.[14] Training of complex surgical tasks is a necessary prerequisite in an advanced surgical training program; therefore, focus needs to be drawn on whether feedback is a requirement when learning these operational skills.

Current literature suggests a predefined proficiency level based on experts' performance as an end point for novice training rather than a fixed training time.[15–18] Having attained a proficiency level on a virtual reality simulator has thoroughly been investigated, and many studies have shown improvement in performance in the operation theater.[2,3,5] These studies all included use of feedback, but their aim was to investigate transfer of skills from the virtual environment to the clinical setting and not how instructor feedback impacts performance during simulator training. At present, it is relevant both to target investigation on how instructor feedback influences learning and to investigate amount of feedback needed to optimize learning. Within motor skills learning, researchers have demonstrated that participants who self-direct their access to instruction or feedback during practice learn more than those whose access is controlled externally,[19,20] but it is uncertain whether these results apply in the surgical training environment.

With a worldwide proliferation of simulation centers, it is essential to explore the optimal circumstances for simulator training and investigate different learning approaches, for example, a

From the *Department of OBGYN, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; †Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark; ‡Department of OBGYN, Hillerød Hospital, Hillerød, Denmark; §Centre of Clinical Education, University of Copenhagen, Copenhagen and Capital Region, Copenhagen, Denmark; ¶Department of Surgery, St. Michael's Hospital, Toronto, Ontario, Canada; and ||Juliane Marie Centre, Centre for Children, Women and Reproduction, Copenhagen, Denmark.

Disclosure: The Copenhagen University Hospital Rigshospitalet and TrygFonden, a nonprofit organization, supported the project. The contents of the manuscript are solely the responsibility of the authors and do not necessarily represent the official views of Rigshospitalet or TrygFonden. The authors declare no conflicts of interest.

The trial complies with the Helsinki Declaration on biomedical research. The Danish National Committee on Biomedical Research Ethics evaluated and approved the trial (journal number: H-3-2010-082). The Danish Data Protection Agency approved collection, analysis, and storage of data (approval code: 2007-58-0015/30-0996).

All participants were provided written information on the trial. Participation was voluntary; no material goods were given to participants. The trial is registered at clinicaltrials.gov with trial registration number: NCT01497782.

Reprints: Jeanett Strandbygaard, PhD, Department of OBGYN, 4074, Rigshospitalet, University Hospital of Copenhagen, Blegdamsvej 9, Copenhagen 2100, Denmark. E-mail: jeanett78@gmail.com.

ISSN: 0003-4932/13/25705-0839
DOI: 10.1097/SLA.0b013e31827eee6e

self-directed approach and the impact and type of feedback. We aimed to investigate the following in a randomized trial: the impact of instructor feedback versus no instructor feedback when training a complex operational task on a virtual reality simulator. Self-directed practice regarding when to receive instructor feedback was applied. In addition, sex differences, computer gaming skills, and self-perception were examined during simulator training.

METHODS

The protocol for this trial has previously been published.[21]

Participants

Medical students in their fourth to sixth years (of 6 years) were recruited through advertisements on Web sites at the surgical and anesthesiological student associations at the Copenhagen University Medical School, Copenhagen, Denmark.[21] The inclusion and exclusion criteria used were as follows:

Inclusion criteria: (1) Medical bachelor degree (completion of the first 3 of 6 years at the University of Copenhagen Medical School). (2) Informed consent before enrolment. (3) Attendance at an introductory meeting before the trial.

Exclusion criteria: (1) Independent experience with more than 3 laparoscopic procedures. (2) Prior experience with virtual reality simulation. (3) Not fluent in the Danish language. (4) Lack of informed consent.

The Virtual Reality Simulator Task and Equipment

The virtual reality simulator was a LapSim, version 2010, produced by Surgical Science, Sweden. The virtual reality task was a right-side laparoscopic salpingectomy due to an ectopic pregnancy. At the end of each completed task, the virtual reality simulator summed up time spent and quality of performance (Table 1) and presented automated feedback (ie, performance score) available to all participants. The virtual reality simulator electronically recorded data from every repetition. These electronic data was transferred to a secure database by an independent investigator.

The Intervention Group and the Control Group

Both the intervention group and the control group had to reach a predefined proficiency level. The predefined proficiency level was defined and validated in a previous study by the same research group[22] and is referred to as the "expert level" in this article. The expert level had to be reached twice within 5 consecutive repetitions.

TABLE 1. Predefined Proficiency Level for the Virtual Reality Simulated Operation Module Right-side Laparoscopic Salpingectomy

Variable	Passing Range	Weight in Calculating Performance Score
Total time	>280 s	15
Blood loss	>180 mL	15
Pool volume	>10 mL	0
Ovary diathermy damage	>3 s	5
Tube cut: uterus distance	>4 mm	5
Bleeding vessel cut	0	Fail if performed
Evacuation from body	>1	Fail if not performed
Left instrument path length	>2 m	15
Left instrument angular path	>350 degrees	15
Right instrument path length	>3 m	15
Right instrument angular path	>450 degrees	15

Based on 11 variables listed, the virtual reality simulator generates a performance score (%) on the operational task: *right-side salpingectomy*, which is available for all participants after each repetition. When all variables are within the passing range, the predefined proficiency level (ie, expert level) is reached. Expert level is set and validated in a previous study.[22] The performance score is calculated on the basis of weighting of the different parameters; the better the individual parameter, the better the performance score.

All participants were informed about the operational technique of a salpingectomy during the obligatory introduction meeting.[21] Furthermore, all participants were instructed on the use of the virtual reality simulator and shown the instruction video both on how to perform the operational task and on how to interpret the automated feedback generated by the simulator before the practice sessions began.

Practice sessions were maximum 3 hours per day; the virtual reality training was not entirely of 3 hours, but it allowed for breaks during training. Participants were instructed to finish the practicing within a 2-month period.

The *intervention group* had an obligatory feedback session oriented toward the first virtual reality operation and could additionally request 2 instructor feedback sessions. One instructor (the first author) provided the standardized feedback and used the same template for every participant. The instructor feedback was 10 to 12 minutes long, and the template-used feedback was standardized and consisted of the following: how to hold the laparoscopic instruments and thereby minimize instrument movements, how to optimally use electric cautery, and how to remove the fallopian tube. The feedback was not tailored to individual needs.

The *control group* did not receive any instructor feedback. Both groups were asked to complete a questionnaire before and after the trial pertaining to perception of own surgical skills on a Likert-type scale ranging from 1 to 5.

Outcome Measures

The outcome measures were as follows:

Primary outcome measures: Number of repetitions and total time (in minutes) used to reach expert level.

Secondary outcome measure: The performance score (%) obtained when expert level was reached. The performance score was automatically generated by the virtual reality simulator and based on performance recorded during the task (Table 1).

In post hoc analyses, the effects of sex and computer gaming skills were explored.

Randomization and Participation

A computer randomization was performed at the Copenhagen Trial Unit. The randomization procedure was concealed and executed by using the participants' unique personal identification number: the central personal registration number. We followed the CONSORT Statement for randomized trials.[21]

Sample Size Calculation

On the basis of data in a previous trial,[4] it was assumed that participants in the intervention group and the control group, on average, would use 30 and 40 repetitions, respectively, to reach expert level (ie, a minimal relevant difference of 10 repetitions). The standard deviation was set to 15. With type I error set at 0.05 and power set at 0.90, the sample size added up to 96 participants, with 48 participants in each group.

Statistical Analysis

The data were analyzed using SPSS (Chicago, IL), version 15.0. Two-sided significance tests were used with a level of significance of 0.05. The distributions of each outcome measure were compared between the intervention group and the control group, using the general linear univariate model. The analyses were repeated, with the covariate semester number and 2 protocol-specified cofactors (sex and computer gaming experience) included.

If the assumptions of the model (normally distributed residuals and variance homogeneity of the groups compared) could not be fulfilled using simple transformations of the data, the distributions of the intervention group and the control group were compared using a nonparametric test (Mann-Whitney). To adjust the P values for multiple testing, the Holm procedure was used.[23]

All of the above-specified analyses were complete participant analyses. If a significant effect of the intervention after adjustment for multiple testing was noted, 3 sensitivity analyses were carried out with a Mann-Whitney test, using mean values of the primary and secondary outcomes: (1) the worst-case scenario (missing value imputed by the most pessimistic value from the opposite group); (2) strong bias (missing value imputed by the most pessimistic value in the group to which it belonged); and (3) mild bias (missing value imputed by the mean value found during the complete participant analysis of the group to which the missing value did not belong). These had increasing degree of bias by replacing missing values with constructed ones reflecting the degree of skepticism of the observed effect.

Figures are presented in error plots showing mean and 95% confidence intervals (CIs). Post hoc analysis examined whether sex and/or computer gaming experience interacted with the intervention. Questionnaire replies concerning perception of own surgical skills after the trial was analyzed with a Mann-Whitney test.

RESULTS

Primary and Secondary Outcome Measures

Participant enrollment and demographics are presented in Figure 1. All participants reached the expert level within a 2-month period. Average time per training session was 2 hours 20 minutes.

Table 2 shows the primary outcomes: the number of repetitions and time (in minutes) to reach the expert level. It was necessary to log transform the data to obtain Gaussian distributions and variance homogeneity. Mean of ln(repetitions) (mean of the control group minus mean of the intervention group) was 0.80 (95% CI, 0.56–1.03; $P < 0.0005$) (Table 2). Mean of ln(minutes) [mean of the control group minus mean of the intervention group was 0.734 (95% CI, 0.544–0.924; $P < 0.0005$)] (Table 2). Repetitions as well as time were significantly higher in the control group than in the intervention group.

The performance score (%) at expert level was significantly lower in the intervention group (mean of control group minus mean of intervention group) 7.23; 95% CI, 2.34–12.1; $P = 0.004$) (Table 2). Similar results were obtained when the analysis was repeated with the protocol-specified cofactors included. Results were still highly significant when adjusted for multiple testing.

Three participants dropped out from the control group in frustration. However, the primary outcomes, repetitions and time, both still differ highly significantly between the 2 groups in all 3 sensitivity analyses ($P = 0.005$ in all 3 analyses). The analysis of the performance score (%) obtained at the expert level was not significant in the worst-case scenario ($P = 0.24$) and strong bias scenario ($P = 0.083$), but significant in mild bias ($P = 0.034$) and original data set scenarios ($P = 0.004$).

Responses from the questionnaire showed that participants' perception of own surgical skills after the trial was significantly higher in the intervention group ($P = 0.011$).

Post Hoc Analysis Regarding Sex and Computer Gaming Skills

Post hoc analysis showed that men in general used significantly less time to reach expert level than women ($P = 0.037$), but no significant difference was observed about the number of repetitions ($P = 0.20$). There was significant interaction between sex and the intervention ($P = 0.044$); the effect of intervention seems to be more pronounced in women than in men (Fig. 2). Borderline significance ($P = 0.051$) was shown about repetitions in favor of men.

No significant interaction was found regarding computer gaming experience and the outcome measures time and repetitions ($P = 0.83$ and $P = 0.88$, respectively). Participants with computer gaming experience had a significantly higher performance score (%) ($P = 0.011$).

Feedback Requests

Two participants did not request the second feedback session; for them, 1 feedback session was sufficient. The second feedback session was requested by 46 participants (96%) and oriented around the tenth repetition (range, 6–11 repetitions). The third feedback session was requested by 17 participants (35%) and was unevenly distributed from the 14th to 50th repetitions. The average time spent on the first, second, and third feedback sessions was 11, 11, and 8 minutes, respectively.

DISCUSSION

Impact of Instructor Feedback

This randomized trial revealed that instructor feedback increases the efficiency regarding the amount of time and number of repetitions needed to reach a predefined proficiency level for a complex simulated operation. The intervention group used less time and fewer repetitions, and results were overall more homogeneous than those obtained for the control group. However, the majority of the control group succeeded in reaching the predefined proficiency level (ie, expert level without instructor feedback). We expected that for this group of novices practicing a complex task, the cognitive load would be overwhelming. This was also the case for 3 participants in the control group who dropped out because of frustration for not being able to reach proficiency level. The control group, however, achieved a significantly higher performance scores than the intervention group, although at a significantly slower pace. This finding was not unexpected; the performance score is a measure based on time and accuracy and increases with training, and in average, the control group used twice the amount of time training.

The control group participants assessed their own surgical skills significantly lower after the trial than the intervention group. The validity of self-assessment is disputable, and there is a difference between confidence and competence; improvement in confidence is not necessarily translated into better competence and better outcome.[24,25] Nonetheless, the point of extensive skills training is not for the trainees to feel a decline of skills, which is an additional argument for supplementary feedback.

Because there were no dropouts from the intervention group, it is likely that feedback would have allowed dropouts from the control group to reach expert level.

Learning of simple and basic laparoscopic tasks, such as coordination and instrument navigation, without instructor provided-augmented feedback has previously been demonstrated in a randomized trial to be more effective than instructor-controlled learning.[14] These basic tasks are relatively easy to accomplish intuitively, which could explain the different findings in our randomized trial where the task was a complex operation involving both knowledge and motor skills. It is feasible that trainees who initially train on the virtual reality simulator could learn simple tasks without instructor feedback. However, on the basis of our findings, when proceeding to more complex tasks, we recommend feedback from an instructor to ensure time-efficient and correct learning.

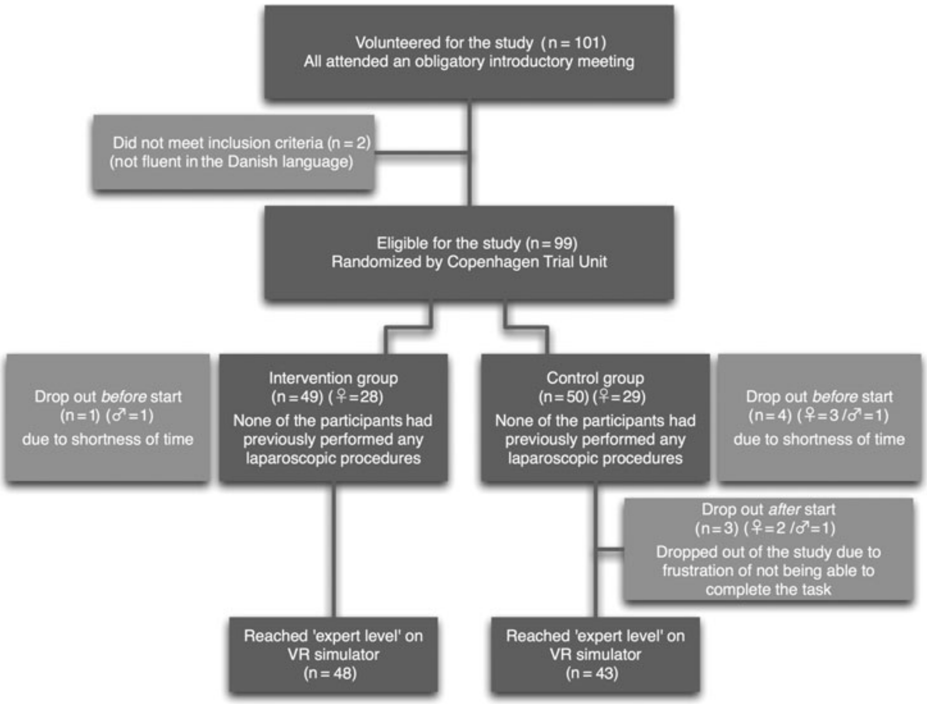

FIGURE 1. Participant enrollment (complying the CONSORT Statement) and demographics. The distribution of sex did not differ significantly between dropouts and those participating in the trial ($P = 0.46$), nor did the presence of computer gaming experience ($P = 0.25$). VR indicates virtual reality.

TABLE 2. Mean Differences in Primary and Secondary Outcome Measures Between the Intervention Group and the Control Group

	Intervention Group (n = 48)	Control Group (n = 43)	*P*
Mean No. repetitions used to reach the expert level	29 (23.9–33.5)	65 (53.9–75.5)	<0.0005
Mean time (min) used to reach the expert level	162 (140–183)	342 (285–398)	<0.0005
Performance score (%) *at* the expert level	49 (45–53)	57 (53–60)	0.004

Values given in parentheses are 95% CI.

Boyle et al[9] found that standardized feedback was associated with significantly fewer errors and improved learning curve when performing a hand-assisted laparoscopic colectomy on a the hybrid simulator ProMIS (a hybrid simulator is a video trainer that provides feedback) among 3- to 5-year surgical trainees. However, they did not report how time consuming or how frequent feedback was provided. Boyle et al[9] argue that feedback could be provided by nonsurgical staff, given the facilitator is sufficiently familiar with the procedure and simulator, which is an argument we endorse but need to explore further. Another randomized trial where the intervention group had access to virtual reality training (but no feedback) and the control group did not train on a virtual reality simulator indicated that simulator training in a nonsupervised setting may not be sufficient to increase laparoscopic suturing skills.[26]

In contrast with previous studies,[27,28] computer gaming experience was not a significant predictor of time and number of repetitions.

No baseline on the virtual reality simulator between the 2 groups was assessed because this familiarity with the virtual reality simulator might have contaminated the outcome. In addition, because of the nature of the randomization and the fact that there was no difference in operational experience between the 2 groups, it is reasonable to believe that the results reflect the true observed differences. All participants received the same standardized information regarding the operational technique for a salpingectomy at the pretrial introductory meeting along with a review of the instruction video generated by the virtual reality simulator. However, whether the 2 groups' understanding of the procedure was the same is unknown; nevertheless, because it is a homogeneous group of participants, it is reasonably fair to state that they had the same minimum knowledge level. We did fully obtain the planned sample size, but few participants had missing values during the trial. The deviation of participants was very small, and we do not believe it affects our results.

Participants indicated that instructor feedback positively affected the trainees' self-perception of surgical skills. This finding is in accordance with other studies focusing on reactions from trainees using simulation-based training. Among the advantages found were improved self-confidence and self-efficacy and improved feeling of being proficient.[29,30] However, there were also some indications of drawbacks, which include high levels of anxiety and stress.[31,32] Little is known about which personal and contextual factors facilitate or impair transfer of learning from a simulation-based setting to the clinical setting, and the next ideal step would be to demonstrate improved clinical performance when having received standardized feedback on a virtual reality simulator. One randomized trial within endoscopic training of general surgery trainees proved better outcome on colonoscopy when systematic feedback on colonoscopy performance was applied.[33] It is unknown whether the skills required for colonoscopy and laparoscopy are the same; nonetheless, it is reasonable to extrapolate the use of systematic feedback to other specialties.

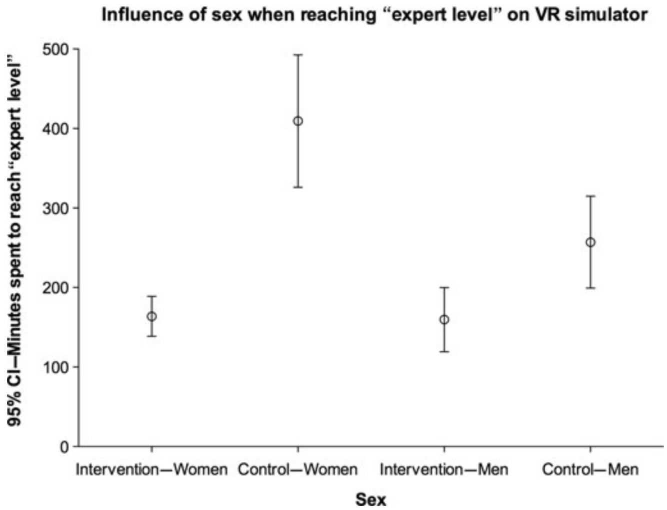

FIGURE 2. Comparison of time used to reach expert level between men and women. The effect of intervention is significantly better in women than in men ($P = 0.044$). Women in the intervention group: mean, 164; 95% CI, 139–189; women in the control group: mean, 409; 95% CI, 326–493. Men in the intervention group: mean, 159; 95% CI, 119–200; men in the control group: mean, 257; 95% CI, 199–315. VR indicates virtual reality.

Sex Differences in Surgical Virtual Reality Training

Men spent significantly less time reaching the predefined proficiency level than women, but no difference was found regarding the number of repetitions. In addition, we found that the feedback influenced the women's performance more than that of men. These findings are in accordance with several other studies, and yet this difference remains largely unrecognized and, consequently, unaddressed.[27,34,35] In an environment where laparoscopic training is increasing, it is important that surgical curricula acknowledge differences between sex to ensure fair and personalized training opportunities.

Predefined Proficiency Level: Is the Bar Set High Enough or Should We "Overtrain"?

Interestingly, we found that the control group reached the expert level with a significantly higher performance score than the intervention group. One interpretation of this is that "practice makes perfect" because, on average, the control group performed twice as many operation modules and used double the amount of time practicing.

Since the introduction of virtual reality simulators in surgical training 10 years ago, the optimal end point for simulation training has been a predefined proficiency level based on that of expert surgeons (ie, surgeons who have performed numerous operations).[15–17] The current standard is that trainees have to accomplish this level once or twice, depending on the set up; in our trial, passing twice was applied.

"Overtraining" of intracorporeal suturing has been shown to be beneficial on the learning curve,[36] and an important determinant of skills and knowledge retention is the amount of "overlearning" or additional training beyond that required for proficiency.[37] Speculations could be made on whether a preset proficiency level based on experts' performance is a sufficient level for training. On the basis of our findings, we suggest that when planning a surgical virtual reality curriculum, an amount of overtraining or continuous training should be considered.

Self-Directed Feedback

A self-directed learner takes responsibility for knowledge production by becoming behaviorally and metacognitively active, and increased autonomy probably allows the participant to tailor knowledge production to his or her specific needs.[19,38] Self-directed access to instruction or feedback has within the sports domain demonstrated higher learning outcomes than that for the participants whose access is controlled externally.[19,20,39] On the basis of this fact, we let the intervention group participants decide themselves when they requested feedback. Contrary to the motor skills research in sports, which showed scattered feedback requests, our results showed feedback requests clustered around the 10th repetition of training. The difference could be explained by the fact that an operation is considered high stakes and calls for feedback sooner, whereas in basketball,[19] it is easier to perform "trial and error" without its perceived consequences. Surprisingly, only one third of the participants wanted the third feedback session; 2 sessions seemed sufficient. To our knowledge, no prior studies have focused on the optimal time to provide feedback in surgical virtual reality training. Our findings with feedback oriented around the first and 10th repetitions could be used as a guideline, although it needs to be examined further whether this result reflects a general tendency in simulator training.

Study Strengths and Limitations

One strength of this randomized trial is the relatively large sample size compared with previous studies on efficiency of virtual reality simulator training. This reduces the risk of random errors.[40] Another strength is the conduct of central randomization and blinded outcome assessment, that is, outcome produced by the virtual reality simulator, which reduces the risk of both selection bias and assessment bias.[40] Instructor feedback was standardized to ensure consistency and reproducibility and can therefore be replicated.

The major limitation is that the sample comprised senior medical students and hence generalizability of results to first-year trainees to whom a virtual reality training curriculum would apply could be a problem. The participating senior medical students had completed all anatomy courses and the mandatory 6-month surgical stay oriented toward the end of medical school, and, in several contexts, the trial participants actually resembled first-year trainees because they often have no prior laparoscopic training either. Moreover, the participants were recruited from special interest groups. In comparison with a previous trial by the same author group where first- and second-year residents performed the same operational virtual reality task (and also received feedback), the average number of repetitions used to reach the predefined proficiency level is almost identical to the average number used by the intervention group.[4] Whether the senior medical students is a true resemblance to first-year residents is unknown, thus we feel confident that motivation to learn this kind of complex technical skills was equally high among the students as in first-year postgraduate trainees.

Future Perspectives

Research within the sports domain has shown that augmented feedback can have a dramatic effect during training. Yet, whether this effect reflects a sustainable capacity to perform the task needs to be demonstrated in retention or transfer studies, for example, performance of a similar but different task or performance in clinical practice on real patients. Both these aspects are topics for future research.

CONCLUSIONS

Instructor feedback increases efficiency when training a complex operational task on a virtual reality simulator; time and repetitions used to achieve a predefined proficiency level were significantly reduced in the group that received instructor feedback compared with the control group.

ACKNOWLEDGMENTS

The authors thank Birgitte Schneider, Rigshospitalet Copenhagen University Hospital, for designing a secure database.

The author contribution is as follows: J.S. (principal investigator) acquired the data, drafted the manuscript, and obtained funding. J.S., T.G., M.M., C.R.L., C.R., J.L.S., B.O., and C.G. contributed to the trial design. P.W., J.S., and C.G. conducted the statistical analyses. F.B. acquired data and monitored trial participants. All authors provided administrative support, analyzed and interpreted the data, and critically revised the manuscript.

REFERENCES

1. Aggarwal R, Ward J, Balasundaram I, et al. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg*. 2007;246:771–779.
2. 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.
3. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg.* 2004;91:146–150.
4. Larsen CR, Soerensen JL, Grantcharov TP, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. *BMJ*. 2009;338:b1802.
5. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg*. 2002;236:458–463.
6. Burden C, Oestergaard J, Larsen CR. Integration of laparoscopic virtual-reality simulation into gynaecology training. *BJOG*. 2011;118(suppl 3):5–10.
7. Stefanidis D, Heniford BT. The formula for a successful laparoscopic skills curriculum. *Arch Surg*. 2009;144:77–82.
8. Schout BMA, Hendrikx AJM, Scheele F, et al. Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surg Endosc.* 2010;24:536–546.
9. Boyle E, Al-Akash M, Gallagher AG, et al. Optimising surgical training: use of feedback to reduce errors during a simulated surgical procedure. *Postgrad Med J.* 2011;87:524–528.
10. O'Connor A, Schwaitzberg SD, Cao CGL. How much feedback is necessary for learning to suture? *Surg Endosc.* 2008;22:1614–1619.
11. Sewell C, Morris D, Blevins NH, et al. Providing metrics and performance feedback in a surgical simulator. *Comput Aided Surg*. 2008;13:63–81.
12. Van Sickle KR, Gallagher AG, Smith CD. The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy. *Surg Endosc*. 2007;21:220–224.
13. Ende J. Feedback in clinical medical education. *JAMA*. 1983;250:777–781.
14. Snyder CW, Vandromme MJ, Tyra SL, et al. Proficiency-based laparoscopic and endoscopic training with virtual reality simulators: a comparison of proctored and independent approaches. *J Surg Educ.* 2009;66: 201–207.
15. Aggarwal R, Grantcharov T, Moorthy K, et al. A competency-based virtual reality training curriculum for the acquisition of laparoscopic psychomotor skill. *Am J Surg*. 2006;191:128–133.
16. Brunner WC, Korndorffer JR, Sierra R, et al. Laparoscopic virtual reality training: are 30 repetitions enough? *J Surg Res*. 2004;122:150–156.
17. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg.* 2005;241:364–372.
18. Brinkman WM, Buzink SN, Alevizos L, et al. Criterion-based laparoscopic training reduces total training time. *Surg Endosc*. 2012;26:1095–1101.
19. Wulf G, Raupach M, Pfeiffer F. Self-controlled observational practice enhances learning. *Res Q Exerc Sport*. 2005;76:107–111.
20. Wulf G, Shea C, Lewthwaite R. Motor skill learning and performance: a review of influential factors. *Med Educ.* 2010;44:75–84.
21. Oestergaard J, Bjerrum F, Maagaard M, et al. Instructor feedback versus no instructor feedback on performance in a laparoscopic virtual reality simulator: a randomized educational trial. *BMC Med Educ.* 2012;12:7.
22. Larsen CR, Grantcharov T, Aggarwal R, et al. Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. *Surg Endosc.* 2006;20:1460–1466.
23. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70.
24. Ward M, Gruppen L, Regehr G. Measuring self-assessment: current state of the art. *Adv Health Sci Educ Theory Pract.* 2002;7:63–80.
25. Duffy FD, Holmboe ES. Self-assessment in lifelong learning and improving performance in practice: physician know thyself. *JAMA*. 2006;296:1137–1139.
26. Halvorsen FH, Fosse E, Mjåland O. Unsupervised virtual reality training may not increase laparoscopic suturing skills. *Surg Laparosc Endosc Percutan Tech*. 2011;21:458–461.
27. Grantcharov TP, Bardram L, Funch-Jensen P, et al. Impact of hand dominance, gender, and experience with computer games on performance in virtual reality laparoscopy. *Surg Endosc*. 2003;17:1082–1085.
28. Van Hove C, Perry KA, Spight DH, et al. Predictors of technical skill acquisition among resident trainees in a laparoscopic skills education program. *World J Surg.* 2008;32:1917–1921.
29. Cass GKS, Crofts JF, Draycott TJ. The use of simulation to teach clinical skills in obstetrics. *Semin Perinatol*. 2011;35:68–73.
30. Zigmont JJ, Kappus LJ, Sudikoff SN. Theoretical foundations of learning through simulation. *Semin Perinatol*. 2011;35:47–51.
31. Bong CL, Lightdale JR, Fredette ME, et al. Effects of simulation versus traditional tutorial-based training on physiologic stress levels among clinicians: a pilot study. *Simul Healthc.* 2010;5:272–278.
32. Harvey A, Nathens AB, Bandiera G, et al. Threat and challenge: cognitive appraisal and stress responses in simulated trauma resuscitations. *Med Educ.* 2010;44:587–594.
33. Harewood GC, Murray F, Winder S, et al. Evaluation of formal feedback on endoscopic competence among trainees: the EFFECT trial. *Ir J Med Sci.* 2008;177:253–256.
34. Elneel FHF, Carter F, Tang B, et al. Extent of innate dexterity and ambidexterity across handedness and gender: implications for training in laparoscopic surgery. *Surg Endosc*. 2008;22:31–37.
35. Thorson CM, Kelly JP, Forse RA, et al. Can we continue to ignore gender differences in performance on simulation trainers? *J Laparoendosc Adv Surg Tech A.* 2011;21:329–333.
36. Kolozsvari NO, Kaneva P, Brace C, et al. Mastery versus the standard proficiency target for basic laparoscopic skill training: effect on skill transfer and retention. *Surg Endosc.* 2011;25:2063–2070.
37. Kneebone RL, Scott W, Darzi A, et al. Simulation and clinical practice: strengthening the relationship. *Med Educ*. 2004;38:1095–1102.
38. Brydges R, Carnahan H, Safir O, et al. How effective is self-guided learning of clinical technical skills? It's all about process. *Med Educ.* 2009;43:507–515.
39. Chiviacowsky S, Wulf G. Self-controlled feedback is effective if it is based on the learner's performance. *Res Q Exerc Sport*. 2005;76:42–48.
40. Keus F, Wetterslev J, Gluud C, et al. Evidence at a glance: error matrix approach for overviewing available evidence. *BMC Med Res Methodol.* 2010;10:90.