The effect of blocked versus random task practice schedules on the acquisition and retention of surgical skills

Justin D. Rivard, M.Sc., M.D.\textsuperscript{a,*}, Ashley S. Vergis, MMedEd, M.D.\textsuperscript{a}, Bertram J. Unger, M.D., Ph.D.\textsuperscript{b}, Lawrence M. Gillman, MMedEd, M.D.\textsuperscript{a}, Krista M. Hardy, M.Sc., M.D.\textsuperscript{a}, Jason Park, MEd, M.D.\textsuperscript{a}

\textsuperscript{a}Department of Surgery and \textsuperscript{b}Medical Education, University of Manitoba, Health Sciences Centre, GF434-820 Sherbrook Street, Winnipeg, Manitoba, R3A 1R9, Canada

Abstract

BACKGROUND: When learning multiple tasks, blocked or random training schedules may be used. We assessed the effects of blocked and random schedules on the acquisition and retention of laparoscopic skills.

METHODS: Thirty-six laparoscopic novices were randomized to practice laparoscopic tasks using blocked, random, or no additional training. Participants performed immediate post-tests, followed by retention tests 6 weeks later. Outcomes included previously validated Fundamentals of Laparoscopic Surgery (FLS) and hand–motion efficiency scores.

RESULTS: Both blocked and random groups had significantly higher FLS and hand–motion efficiency scores over baseline on post-tests for each task (\(P < .05\)) and higher overall FLS scores than controls on retention tests (\(P < .01\)). No difference was seen between the blocked and random groups in the amount of skill acquired or skill retained.

CONCLUSIONS: Both blocked and random training schedules can be considered as valid training options to allow programs and learners to tailor training to their individual needs.

Technical skills are an essential component of a surgical residency that have traditionally been taught in the operating room. However, an increasing number of surgical residency programs have now incorporated basic skill laboratories and simulation into their training, with the goal of increasing operating room efficiency, operating room experience, and patient safety.\textsuperscript{1-4} Despite the proliferation of simulator use in residency programs, relatively little is known about the best way to implement simulation into surgical training to promote long-term learning.

To teach and assess basic laparoscopic skills, the American College of Surgeons and the Society of American Gastrointestinal and Endoscopic Surgeons have endorsed the bench model simulator-based Fundamentals of Laparoscopic Surgery (FLS) program.\textsuperscript{5} Despite the popularity and widespread implementation of the FLS program, there has been little systemic study on how the FLS program and its...
individual components are best applied within the context of broader technical skills curricula. A program or course that consists of multiple discrete tasks can be practiced in a blocked or random fashion. A blocked practice schedule is one in which a single task is practiced repeatedly before moving on to the next task. In contrast, random practice schedules involve performing the same number of repetitions of each task but in a random order such that a given task is never practiced on successive trials. Published FLS training protocols instructed learners to practice the 5 FLS tasks in a blocked fashion. 6,7 It is unclear, however, whether the FLS-blocked training protocol is optimal for learning laparoscopic skills.

In the motor learning literature, Shea and Morgan 8 previously assessed learning simple motor skills in a blocked and random fashion. On the immediate post-test, there was a clear advantage for the blocked practice group. However, on the delayed retention tests given 10 days later, there was an advantage for the random practice group. These findings refute the assumption that the condition that speeds skill acquisition is also most effective for skill retention. Repeated practice in a blocked fashion may help participants acquire motor skills more quickly, but these skills may not necessarily be retained over time. Conversely, participants who follow random practice schedules may acquire skills more slowly but may be better equipped to retain their skill over time. Although studies in the motor learning literature demonstrate an advantage for long-term learning under random practice conditions, the applicability of these findings to surgical skills training is not entirely clear. 9

The goal of any training program is to convey long-term learning on the learners. This is especially salient in a surgical residency where it is rare to practice a task and then immediately have the opportunity to perform it in an operative setting. This study examines the effects of blocked and random schedules of task practice on the acquisition and retention of laparoscopic skills.

Methods

Participants

General surgery, urology, plastic surgery, orthopedic surgery, otolaryngology, cardiac surgery, and obstetrics and gynecology residents in postgraduate years (PGYs) 1 to 2 and senior medical students with an interest in a surgical specialty attending the University of Manitoba were invited to participate in this study. Individuals who had previously completed the FLS course or had previous experience performing laparoscopic surgery (defined as the primary surgeon for any laparoscopic procedure) were excluded. An effect size of 1.2 standard deviations (SDs) was considered a large but acceptable difference in assessing the teaching intervention in this study. This is consistent with the surgical skills and psychology training literature. 10–12 To detect this effect size using a 2-tailed alpha of .05 and a power of .8, 12 subjects were required in each practice schedule arm. Approval was obtained from the University of Manitoba Health Research Ethics Board.

Study design

Skill practice and testing took place using the FLS box trainer (VTI Medical, Waltham, MA) between October 2011 and March 2012. All participants were oriented to the box trainer and watched a video tutorial (FLS CD ROM Disk #2). Participants then performed 2 repetitions of 4 FLS tasks (peg transfer, pattern cutting, ligating loop, intracorporeal [IC] suturing), as previously described. 13 The mean of these 2 trials constituted their baseline pretest scores. All individuals were then randomized into 1 of 3 groups using unmarked, sealed envelopes: (1) no additional training (control group), (2) FLS training using a blocked task schedule, or (3) FLS training using a random task schedule. Medical students and residents were stratified separately. Participants randomized to either training group performed the following number of repetitions for each task in their practice sessions: 12 peg transfers, 6 pattern cuts, 4 ligating loop applications, and 10 IC sutures. Previous studies suggest that participants will show measurable improvements in task performance with these repetition numbers, although they were very unlikely to achieve expert levels of proficiency. 6,14–16 Participants in the blocked schedule group performed all repetitions of a single task before moving on to the next task. Participants in the random schedule group performed each task in a random order, as predetermined by the authors, with no task being repeated on 2 consecutive trials. The total number of repetitions was constant between experimental arms and only the order of task practice varied, with all subjects within each group performing the same order of tasks.

All participants in the blocked and random practice groups performed a post-test immediately after training as a measure of skill acquisition. All 3 groups then performed retention tests 4 to 6 weeks after training. The retention tests consisted of performing 2 sets of the 4 tasks, once in the blocked order and once in the random order. To minimize the specificity of learning caused from the practice conditions, 17 half the participants in each experimental group performed the test tasks first in a blocked and then in a random order, whereas the other half performed the test tasks first in a random followed by a blocked order.

Outcome measures

Previously validated outcome measures on each of the tests included: (1) FLS scoring metrics 13,16,18 and (2) computer-based evaluations. 19–24 In the FLS scoring system, a cutoff time is assigned for each task. 13,16 A raw score was then calculated by subtracting the time taken to complete the task (in seconds) and any additional penalty points from the cutoff time, as previously described. 15 Each individual task has predetermined penalties associated with its own penalty score.

The computer-based performance measures consisted of completion time, number of hand movements, and path
length traveled. Hand motion data were collected with an electromagnetic tracking system (Ascencion Technology Corporation, Burlington, VT) and analyzed using a customized software program (Motion Analysis and Recording System [MARS]; University of Manitoba). This system tracks and records the position of sensors (6° of freedom) placed on the dorsum of the participants’ hands. For motion-tracking purposes, a movement is defined as a change in velocity. MARS uses a Gaussian filter with a width of 12 samples per SD, which corresponds to a high-frequency cutoff filter of 1.666 Hz when the sample rate is 20 Hz. This eliminates unwanted high-frequency background noise (such as hand tremors), ensuring that only meaningful actions are recorded and reported. A velocity threshold of 15 mm/s was used to report an actual movement was instituted. These settings were chosen based on previous calibration experiments reported in the literature. The data collected from the left and right hands were added together to give the total number of movements and path length traveled for each task.

Skill acquisition was measured by subtracting the pretest from the post-test scores for both the FLS and computer-based performance measures. Higher scores indicated a higher level of skill acquisition. Skill retention was calculated by subtracting the post-test score from the retention score. Additionally, skill learning was calculated by subtracting the pretest score from the retention score. Again, higher scores indicated more skill retention and more skill learning.

### Statistical analysis

The FLS scores were normally distributed. Separate 1-way analysis of variance and 1-way analysis of covariance with pretest scores as the covariate were used to analyze the FLS task scores, the overall scores, acquisition scores, and retention scores. Paired t tests were used to assess the amount of skill acquisition and retention that occurred within each group. Most of the variables in the computer-based hand–motion data (time, path length, number of movements) were assessed to be in a non-normal distribution, and therefore, Kruskal–Wallis tests were used to compare the 3 groups. Significant differences at $P < .05$ were further assessed using the Mann–Whitney U test with Bonferroni correction, giving us a significance level of $P < .017$. Paired data were assessed using the Wilcoxon signed rank test.

### Results

#### Participant characteristics

A total of 36 participants (14 male and 22 female) volunteered for this study (see Table 1 for the baseline characteristics). Participants included senior medical students ($n = 10$), PGY-1 ($n = 17$), and PGY-2 ($n = 9$). There was no significant difference among the 3 study groups’ characteristics at baseline or the average time between the training sessions and retention test.

### FLS scores

Pretest FLS scores were similar for the blocked, random, and control groups (data not shown). On the completion of training for both the blocked and random groups, there was a statistically significant improvement in skill acquisition from pretest to post-test scores (Fig. 1, $P < .01$ for improvement in all scores). Improvements in performance from the pretest to the post-test scores were not significantly different between the blocked and random groups for any of the tasks.

Skill retention tests were performed between 4 and 6 weeks after the initial training. For the blocked training group, a significant decline in performance from post-test scores was seen for the peg transfer, ligating loop, and IC suturing tasks (mean score [SD], peg: post-test $= 93$ [7.3], retention $= 87.2$ [6.0], $P < .01$; pattern: post-test $= 56.1$ [23.8], retention $= 40.1$ [29.0], $P = .12$; loop: post-test $= 71.9$ [17.2], retention $= 55.9$ [15.6], $P = .02$; suturing: post-test $= 78.8$ [8.3], retention $= 60.1$ [23.2], $P = .01$).

For the random training group, a significant decline in performance from post-test scores was seen for the peg transfer, pattern cutting, and ligating loop tasks (mean score [SD], peg: post-test $= 94.7$ [5.3], retention $= 87.4$ [12.7], $P = .01$; pattern: post-test $= 63.4$ [18.2], retention $= 34.5$ [24.9], $P = .01$; loop: post-test $= 77.7$ [12.2], retention...
Figure 1 FLS performance score improvement from pretest to post-test for the blocked and random groups for all tasks. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

 amounted to learning, as defined by the difference between the pretest score and retention score, is shown for each of the 3 groups in Fig. 2. When the groups were divided into any training compared with no training, more skill learning occurred in the training groups (P < .05 for all tasks). However, there were again no differences seen between the blocked and random training groups.

Hand motion analysis

There were no pretest differences seen between the 3 groups for the number of movements, path length traveled, or time taken to complete the task (data not shown). Similar to the FLS scores, both the blocked and random groups showed a significant improvement in performance between pretest and post-test with less movements required, less total path length traveled, and less time required to complete the task (see Fig. 3, P < .05 for all tasks). Improvements in post-test scores were similar between the blocked and random groups for the number of hand movements and path length traveled for all tasks. The only difference between the blocked and random groups was for the pattern cutting task, with the random group completing the task faster (blocked = 276, interquartile range [IQR] 202 to 293 seconds; random = 213, IQR 177 to 246 seconds, P < .05; Fig. 3).

After completion of the skill retention tests, the blocked group showed a significant increase in the median number of hand movements only for the suturing task (IQR), suturing: post-test = 205 (176 to 254), retention = 277 (199 to 353), P = .02. The blocked group also showed an increased median path length traveled for the ligating loop and suturing tasks (meters traveled [IQR], loop: post-test = 10.3 [8.5 to 11.6], retention = 13.1 [12.1 to 15.2], P = .01; suturing: post-test = 14.3 [12.3 to 15.6], retention = 19.5 [16.5 to 22.8], P = .01). The median total time to complete the task for the blocked group significantly increased for the peg transfer, ligating loop, and suturing tasks (time in seconds [IQR], peg: post-test = 80 [74 to 87], retention = 98 [93 to 109], P = .03; loop: post-test = 68 [59 to 100], retention = 95 [79 to 121], P = .03; suturing: post-test = 192 [183 to 217], retention = 254 [202 to 346], P = .01). No difference was seen in the other tasks.

For the random group, a significant increase in median number, path length, and time was seen for the peg transfer and ligating loop tasks (median number of hand movements required [IQR], peg: post-test = 112 [96 to 123], retention = 127 [107 to 152], P = .01, and loop: post-test = 63 [55 to 84], retention = 129 [72 to 173], P < .01; median path length traveled in meters [IQR], peg: post-test = 5.6 [4.8 to 6.9], retention = 6.6 [5.8 to 8.0], P < .01, and loop: post-test = 9.1 [8.5 to 10.1], retention = 12.6 [11.2 to 16.8], P < .01); and median time needed to complete the tasks in seconds [IQR], peg: post-test = 75 [66 to 81], retention = 86 [72 to 114], P < .01, and loop: post-test = 63 [58 to 75], retention = 116 [87 to 143], P < .01). Additionally, the random group also showed a significant increase in median time to complete the task for the pattern cutting task (IQR) pattern: post-test = 213 (177 to 246), retention = 273 (232 to 358), P = .01. No difference was seen in the other tasks.

When the blocked and random groups were combined and compared with the no training group, more skill learning occurred (P < .05 for all tasks). The amount of skill learning
for the hand motion analysis scores is shown in Fig. 4. There were no differences seen between the blocked and random training groups.

**Comments**

The varying of practice schedules into blocked or random fashion has been termed contextual interference (CI) in the motor learning literature. The CI effect suggests that high CI (random practice) slows the speed of acquisition but allows for greater retention of motor skills and that low CI (blocked practice) allows for faster acquisition of motor skills but may be detrimental to long-term learning. Two main theories have been brought forward to help explain this effect. The first is the “action–plan reconstruction” hypothesis. This hypothesis states that before a movement occurs, an “action plan” must be prepared. In blocked practice, the action plan is readily available but it had lack of attention on trials after initial retrieval from working memory. Random learners are forced to repeatedly undergo a demanding reconstructive process to replan the way in which they perform each task. It is this need to repeatedly plan the movement solution that results in poorer performance during acquisition. Presumably, this additional trial-to-trial preparation used by the random group results in a more resilient memory representation that better supports long-term skill recall.

The second theory is the “elaboration–distinctiveness view,” which states that during random practice all tasks remain in working memory, allowing the learner many opportunities to compare and contrast the tasks. The need to keep the tasks separate during practice is what causes the disadvantage during acquisition. During
blocked practice, repeating a long series of the same task makes it less important to keep track of which task is which. Although more demanding during acquisition, the need to compare and contrast yields superior performance in retention tests. In the action–plan reconstruction hypothesis, the previous task is “dumped” from working memory before starting the next task and then regenerated the next time the task needs to be performed. Whereas in the elaboration–distinctiveness view, the tasks all reside beside each other at the same time in working memory during practice. Both theories come to the same conclusion by explaining opposite affects of working memory.

We found that novice trainees learned and retained some laparoscopic skills using either blocked or random practice schedules, but we were unable to demonstrate a clear advantage for one schedule over the other. Two previous studies in the literature of the surgical skills examined blocked and random practice schedules using separate bone-plating tasks. Neither Dubrowski et al. nor Brydges et al. found a difference in long-term learning outcomes between the random and blocked groups.

Thus, although there is a substantial body of evidence that support an advantage for long-term learning under random practice conditions, there is little evidence to support the long-term advantages for random schedules in the surgical literature. There are several possible reasons for these contrasting results. The first is that surgery is more complex and may require more cognitive involvement than other motor skills. Albaret and Thon studied the effect of task complexity on CI and found that the effects of CI were only present on simpler tasks. They hypothesized that if movements are complex, subjects would have difficulty maintaining all movement-related information in working memory between trials. This would allow for more

Figure 4  Individual task skill learning for each of the groups using HMA scores for (A) number of movements, (B) distance traveled, and (C) time to complete the task. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)
complex processing strategies even for the blocked group and would, therefore, hide the beneficial effect of random practice. The FLS tasks used in our study are considerably more complex than many of the motor tasks shown to produce the CI effect in the literature. A second possible reason is that our study design may not have allowed enough trials for the effects of CI to occur. Tsutusi et al studied the CI effect using complex motor tasks and found that it took 15 trials before a difference in acquisition was seen between the blocked and random groups. This suggests that for some skills, practice schedules may not influence the processing operations that are involved in figuring out what to do but instead may manifest more during later stages of learning. A third possible reason for the contrasting results may have been that too much time passed before conducting the retention tests.

The effects of practice conditions in motor learning within the constraints of skill level and task difficulty were brought together by Guadagnoli and Lee into the Challenge Point Framework. This framework suggests that task difficulties create a learning potential whose function differs according to the level of the performer, complexity of the task, and the training environment. For a simple task, a basic action plan may be developed within the first couple attempts, and further refinement of the skill will be dependent on the extent to which the learner is challenged by the practice conditions. A more complex task may require considerably more time to learn and would require more effort and information processing activities on the part of the learner. Introducing additional demands for the learner during this process could in fact be detrimental because the additional demands may be competing for a limited amount of processing capacity. This same relationship may exist for experienced vs novice surgeons. When a surgeon is in a later stage of learning, the ability of the systems to process information improves, and thus, the learner can and should handle more demanding acquisition protocols. The more complex a task is, the more basic (blocked) practice should take place before introducing any additional constraints on the learner such as random practice schedules. Early practice should focus on getting it right. Only once a learner knows what to do, will the true benefit from a random practice schedule appear.

This study has several significant strengths, including a theoretical basis in the motor learning literature, a randomized controlled design, and the use of multiple validated outcome measures. A methodologic limitation that requires mention, however, relates to our sample size calculations. The effect size that we used may have been adequate to discriminate between training and no training groups, but the effect size seen between 2 different training groups may be smaller. Assuming an effect size of .6 would be sufficient to show a difference between the blocked and random groups, 45 subjects per group would have been required. Thus, although we were able to show that both training interventions were beneficial over the control group, more power may be required to detect a difference between the blocked and random practice groups.

In conclusion, a benefit for learning basic laparoscopic skills with both the blocked and random formats was demonstrated. However, an advantage for one format over the other was not determined. Current published FLS training protocols endorse blocked training. This study suggests that both training schedules can be considered options for learners. Furthermore, there is strong evidence in the motor learning literature to support random practice formats, and further study is required to better understand how this applies to surgical skills education.

References