Self-directed practice schedule enhances learning of suturing skills

Oleg Safir, MD, MEd*
Camille K. Williams, MHSc†‡
Adam Dubrowski, PhD§
David Backstein, MD, MEd*
Heather Carnahan, PhD¶

From the *Department of Surgery, †Wilson Centre for Research in Education and ‡Graduate Department of Rehabilitation Science, University of Toronto, the §Disciplines of Emergency Medicine and Pediatrics, Faculty of Medicine, Memorial University of Newfoundland, and the ¶School of Human Kinetics and Recreation, Memorial University of Newfoundland, St. John's, NL

Portions of this work have been previously presented at the Annual Meeting of the Royal College of Physicians and Surgeons of Canada in Ottawa, Ont., September 2008, and the Annual Meeting of the Association for Surgical Education in Salt Lake City, Utah, Apr. 28–May 2, 2009.

Accepted for publication
Jan. 9, 2013

Correspondence to:
H. Carnahan
School of Human Kinetics and Recreation
Memorial University of Newfoundland
St. John's NL A1C 5S7
hcarnahan@mun.ca

DOI: 10.1503/cjs.019512

Background: Most preoperative surgical training programs experience challenges with the availability of expert surgeons to teach trainees. Some research suggests that trainees may benefit from being allowed to actively shape their learning environments, which could alleviate some of the time and resource pressures in surgical training. The purpose of this study was to investigate the effects of self-directed or prescribed practice schedules (random or blocked) on learning suturing skills.

Methods: Participants watched an instructional video for simple interrupted, vertical mattress and horizontal mattress suturing then completed a pretest to assess baseline skills. Participants were assigned to 1 of 4 practice groups: self-directed practice schedule, prescribed blocked practice schedule, prescribed random practice schedule or matched to the self-directed group (control). Practice of the skill was followed by a delayed (1 h) posttest. Improvement from pretest to posttest was determined based on differences in performance time and expert-based assessments.

Results: Analyses revealed a significant effect of group for difference in performance time of the simple interrupted suture. Random practice did not show the expected advantage for skill learning, but there was an advantage of self-directed practice.

Conclusion: Self-directed practice schedules may be desirable for optimal learning of simple technical skills, even when expert instruction is available. Instructors must also take into account the interaction between task difficulty and conditions of practice to develop ideal training environments.

Contexte : La plupart des programmes de formation préopératoire en chirurgie ont du mal à trouver des experts pour enseigner la chirurgie aux stagiaires. Selon certaines recherches, il pourrait être utile de permettre aux stagiaires de structurer eux-mêmes leurs milieux d'apprentissage, ce qui pourrait se révéler avantageux compte tenu des contraintes de temps et de ressources. Le but de cette étude était de mesurer les effets de différents horaires d'exercices, autodirigés ou prescrits (aléatoires ou fixes), sur l'apprentissage des techniques de sutures.

Méthodes : Les participants ont regardé une vidéo de formation sur les sutures uniques interrompues de type matelassier verticales et horizontales, avant de subir un prétest pour évaluer leurs compétences de base. Les participants ont ensuite été assignés à 1 de 4 groupes de pratique : horaires d'exercices autodirigés, fixes prescrits, aléatoires prescrits ou assortis au groupe « autodirigé » (témoin). La période d'exercices était suivie d'un post-test administré après un délai d'une heure. L'amélioration des résultats entre le prétest et le post-test a été déterminée par les différences de temps d'exécution et l'évaluation d'un expert.

Résultats : Les analyses ont révélé un effet significatif selon le groupe en ce qui a trait aux différences de temps d'exécution pour la suture simple interrompue. Les horaires d'exercices aléatoires ne se sont pas accompagnés de l'avantage attendu pour ce qui est de l'apprentissage de la technique, mais on a noté un avantage associé aux horaires d'exercices autodirigés.

Conclusion : Les horaires d'exercices autodirigés pourraient être souhaitables pour un apprentissage optimal des habiletés techniques simples, même lorsqu'un enseignement par les experts est disponible. Les instructeurs doivent aussi tenir compte de l'interaction entre la difficulté de la tâche et les conditions dans lesquelles se font les exercices pour améliorer les milieux de formation.

Surgical instructors often predetermine the details of training sessions, such as the order of practised tasks, the duration of practice, and timing and type of feedback, while the trainee remains relatively passive. Such training environments are often not tailored to the informational needs of individual learners and, as such, are likely to be suboptimal learning environments. However, as surgical training programs contend with limited instructor availability and time allotted to teach fundamental technical skills, there is a trend toward the use of self-directed learning modules (e.g., CD-ROM and online programs) that require learners to be more active and independent. Recent research suggests that self-directed practice might assist surgical educators in creating learning environments that better support trainees' motivation to practise by meeting changing informational needs.[1] A study conducted by Jowett and colleagues[2] demonstrated that skill performance for tying knots was unaffected by practice enforced after trainees decided they had reached proficiency and did not require further practice. These researchers speculated that the learning environment (e.g., simulation model and instructions) remained unchanged after the trainees reached a certain proficiency on the skill and therefore did not provide any additional benefit to learning. Studies using laboratory, sporting and surgical tasks have shown that motor learning can be facilitated if the learner is able to self-direct various aspects of their training experience, such as frequency of feedback,[3–6] access to video instruction[6,7] and order of practised tasks.[1]

Surgical educators have also begun to consider how other aspects of the training environment, such as distribution and schedule of practice, can be optimized to enhance motor learning within time constraints.[8] With respect to practice schedules, the literature has shown that performance during acquisition of related tasks practised in a random or unsystematic order (e.g., ACB, BAC, ABC, for 3 tasks A, B and C) is impaired compared with performance of tasks in blocked or drill-type order (e.g., AAA, BBB, CCC). Interestingly, however, after a rest period, performance is better on a delayed test for random practice compared with blocked practice. This phenomenon is referred to as the contextual interference effect[9] and is often explained by the forgetting hypothesis,[10] which suggests that practice that forces the learner to repetitively forget and recall the required skills, such as random practice, will enhance delayed performance. This is important because, unlike immediate postpractice performance, which often represents transient practice effects, delayed tests are more likely to reflect relatively permanent improvement in ability (i.e., learning).[11] Since the motor learning literature has shown the contextual interference effect to be greatest for simple laboratory tasks,[12,13] researchers have begun to explore its applicability to more complex tasks in the surgical domain. While these studies have shown divergent results,[14–16] they highlight the fact that, in the surgical domain, interactions between performance and practice variability can be affected by task difficulty.

While there is some evidence supporting random and self-directed practice in the surgical domain, to our knowledge, no studies have examined both practice schedule and the selection of practice schedule for the same surgical skill. The purpose of this study was to determine how practice schedule (random or blocked) and selection of practice schedule (self-directed or prescribed) contribute to learning of suturing skills. We hypothesized that self-directed and random practice schedules would produce better postpractice performance than the other prescribed practice schedules.

Methods

Participants

We recruited first- and second-year medical students from the University of Toronto to participate in our study. The University of Toronto and Mount Sinai Hospital Research Ethics Boards approved the research protocol, and all participants provided voluntary informed consent.

Procedure

Each participant viewed an 8-minute instructional video of an expert demonstration of 3 types of wound closure skills: simple interrupted, vertical mattress and horizontal mattress. Using Sofsilk 3–0 silk sutures (United States Surgical Corporation, Covidien), a synthetic skin pad (Limbs & Things), curved needle, forceps and a needle driver, all participants performed a pretest consisting of 1 trial of each of the 3 sutures without any feedback or access to the instructional video. Each participant was then randomly assigned to 1 of 4 practice schedule groups: self-directed, random, blocked or matched control. Participants in the self-directed group were free to choose their practice schedules with the constraint that by the end of practice they had performed 5 trials of each suture type. The random and blocked groups received prescribed practice protocols and practised the 3 suture types in the orders outlined in Table 1. Participants in the matched control group served as a control, such that each participant in this group was prescribed exactly the same practice schedule as 1 participant from the self-directed group. The difference between the random and self-directed groups (other than the prescription of the practice schedule) is that the random group had a truly random schedule whereas the self-directed group may have selected elements of both blocked and random patterns in their schedules. During practice, all participants were free to review the instructional video as frequently as they wished. After the practice session there was a rest interval of 1 hour followed by a posttest administered in the same manner as the pretest.

Statistical analysis

The pretest and posttest performances were videotaped and used to obtain measures of performance to assess learning. First, total time to complete each suture (performance time), from the first needle puncture in the skin pad to cutting of the final sutures, was extracted from each video. Second, the pretest videos were independently assessed by 2 experts blinded to the experimental condition or group. The expert observers used 3 validated measures to assess performance across all 3 suture techniques: a global rating scale of operative performance[17] (maximum score 35), a checklist for suture of skin laceration[17] (maximum score 11) and a final product analysis[18] (maximum score 4). These scores were used to assess interrater agreement and consistency by calculating single-measures intraclass correlation coefficients (ICCs) with 95% confidence intervals (CIs) using a 2-way, random-effects model with both absolute agreement and consistency methods. Absolute agreement implies that the raters assigned similar scores (absolute values) for similar performances, whereas consistency means that the raters' scores followed similar trends for the performances even if the absolute scores were not the same. For formative or summative classroom-type assessment, reliability is expected to be in the range of 0.70–0.79 or lower depending on the length of the test/number of test items.[19]

One rater then went on to score the posttest videos using the same measures. Difference scores were calculated for performance time and each of the validated measures by subtracting the pretest score from the posttest score. The difference scores were analyzed in separate 4-group 1-way analyses of variance. We considered effects to be significant at $p < 0.05$, and they were further analyzed using the Tukey honestly significant difference method for post hoc comparison of means. We also calculated Cohen *d* effect sizes (using average standard deviations) to help determine the importance of group effects independent of sample size. Effect sizes at 0.2, 0.5 and 0.8 were considered small, medium and large, respectively.

Results

Thirty-eight first- and second-year medical students (20 women and 18 men with a mean age of 23 yr) from the University of Toronto participated in our study. Ten participants were randomly assigned to the self-directed group, 9 to the random group, 10 to the blocked group and 9 to the matched control group. Inspection of the practice schedules selected by the participants in the self-directed group showed that 7 of 10 participants chose a blocked practice schedule. The order of the blocks was the same as the order of suture types demonstrated on the instructional video. The remaining 3 participants chose a hybrid schedule that was predominantly blocked, but had some elements of randomization that appeared later in practice.

Interrater consistency and agreement

Intraclass correlations were calculated for the 2 raters who viewed the pretest videos to determine whether scores from 2 independent raters were consistent and/or in agreement. As seen in Table 2, the ICCs indicated that there was moderate agreement and consistency between the raters for each measurement tool (global rating scale, checklist, final product analysis). The global rating scale had the highest agreement and consistency, followed by the checklist and the final product analysis. However, overlap among the CIs suggests that there were no significant differences in agreement or consistency among the measurement

Table 1. Practice schedules used by participants in each group

			Self-directed (P1–P10) and matched control (P1–P9)									
Trial	Random	Blocked	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
1	V	S	S	S	S	S	S	S	S	S	S	S
2	S	S	S	S	S	S	S	S	S	S	S	S
3	H	S	S	S	S	S	S	S	S	S	S	S
4	H	S	S	S	S	S	S	S	S	S	S	S
5	V	S	V	S	S	S	S	S	S	S	S	V
6	S	V	V	V	V	V	V	V	V	V	V	V
7	V	V	V	V	V	V	V	V	V	V	V	V
8	H	V	V	V	V	V	V	V	V	V	H	V
9	S	V	H	V	V	V	V	V	V	V	H	H
10	V	V	H	V	V	V	V	V	V	V	V	H
11	H	H	H	H	H	H	H	H	H	H	V	H
12	S	H	H	H	H	H	H	H	H	H	V	H
13	V	H	S	H	H	H	H	H	H	H	H	S
14	H	H	V	H	H	H	H	H	H	H	H	V
15	S	H	H	H	H	H	H	H	H	H	H	H

H = horizontal mattress suture; S = simple interrupted suture; V = vertical mattress suture.

tools. The ICCs were all in the range 0.50–0.70, which we believe is acceptable considering the number of items used for each measurement tool. Since we determined that the scores were fairly stable across raters, only 1 rater continued evaluation of the posttest videos.

Performance time

Means and standard errors for differences in performance time from pretest to posttest are shown in Figure 1. The analysis of performance time showed a main effect of group for the simple suture ($F_{3,27} = 4.04$, $p = 0.017$). Tukey post hoc comparisons indicated that self-directed participants decreased their performance time significantly more than both the blocked ($p = 0.036$) and matched control ($p = 0.040$) groups. Further, the effect sizes for both these comparisons ($d = 1.2$ and $d = 1.3$, respectively, exceeded Cohen's convention for a large effect. Performance times for the other 2 suture techniques (horizontal and vertical mattress) showed similar patterns, but these effects were not significant.

Expert ratings

The analyses of the expert ratings did not show any significant effects of group. However, as seen in Figure 2, the trends suggest that on average the self-directed group demonstrated greater or similar improvements from pretest to posttest, particularly using the global rating scale. These results are similar to those for performance time of the simple suture. In fact, since these trends were similar, we calculated Cohen *d* effect sizes to compare each group to the self-directed group. The effect sizes were $d = 1.1$, $d = 1.3$ and $d = 0.8$ for comparisons of the self-directed group with the random, blocked and matched control groups, respectively.

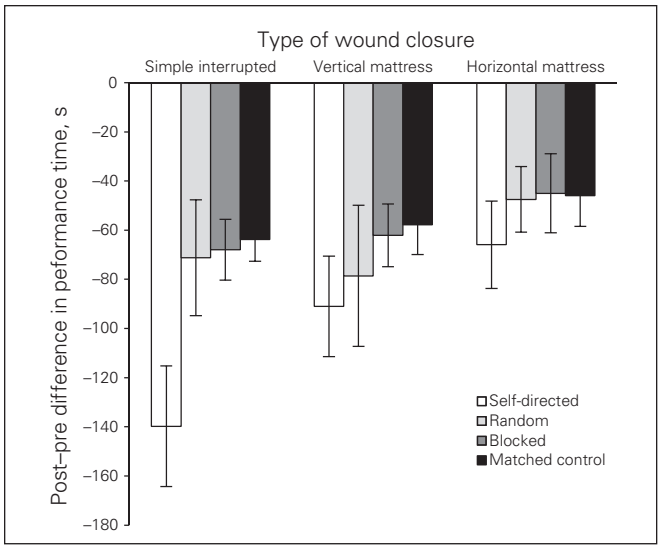

Fig. 1. Comparison of means (with standard errors) for differences in performance time (posttest – pretest) for each practice group performing each type of suture.

Discussion

Our results demonstrate that when compared with the blocked and matched control groups, the self-directed group experienced a significant improvement in performance time for the simple interrupted suture. Furthermore, both these effect sizes were large, suggesting that these effects may have practical importance for training of suturing skills. Our analysis of expert evaluations using a global rating scale, checklist and final product analysis yielded no significant group effects, but trends were similar to those observed for performance time. In addition, effect size calculations for the global ratings showed large effects for comparisons between the self-directed group and all other practice groups, suggesting that these results are also important for future research in this area.

Unlike most studies in the basic motor control literature but in line with the surgical training literature, our results do not support the contextual interference effect; that is, they do not show a learning advantage for random compared with blocked practice. This suggests that random

Table 2. Summary of intraclass correlation coefficients* indicating agreement and consistency between 2 expert raters for pretest scores

	Method; ICC (95% CI)	
Measurement tool	Absolute agreement	Consistency
Global rating scale	0.64 (0.20–0.84)	0.73 (0.51–0.87)
Checklist	0.55 (0.21–0.77)	0.61 (0.32–0.80)
Final product analysis	0.50 (0.20–0.71)	0.55 (0.27–0.74)

CI = confidence interval; ICC = intraclass correlation coefficient.
*Single measures ICC based on 2 raters; 95% CI for estimate of ICC.

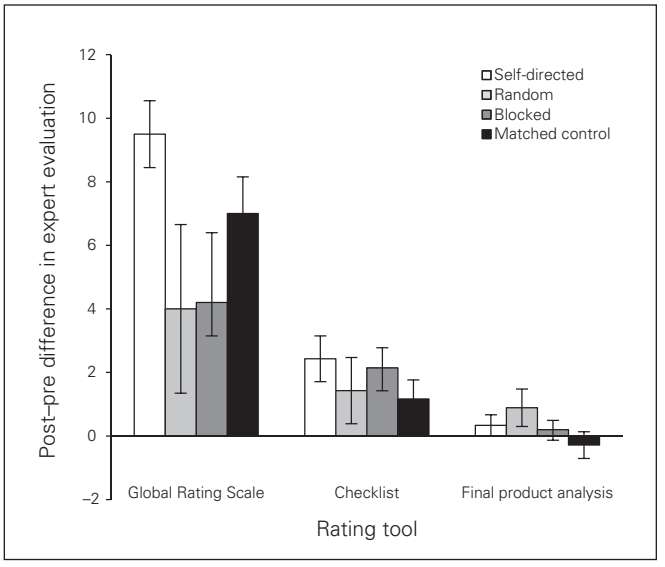

Fig. 2. Comparison of means (with standard errors) for differences in expert evaluation scores (posttest – pretest) using the global rating scale, checklist and final product analysis for each practice group.

practice does not facilitate improved performance for this particular skill (suturing) and adds support to the idea that random practice does not always confer an advantage for learning, particularly for more complex skills like those often explored in the surgical domain.[2] One may argue that the 3 suture types used were quite similar, with only minor variations in the details of the general suturing task. It is then possible that these small variations may not be enough to force the learner to forget and recall the skills; hence, not enough to produce the contextual interference effect. However, in a recent review, Merbah and Meulemans[20] conclude that for more complex applied tasks, the contextual interference effect can appear even when there are variations only within the same type of task. This notion of task complexity and its interactions with motor learning effects should always be considered and is addressed in the challenge point framework, described by Guadagnoli and Lee.[11] They proposed that the effectiveness of a particular practice condition depends on an interaction between the difficulty of the task and the expertise of the learner. It is possible that, for the medical students who participated in the present study, the psychomotor demands of learning 3 suturing techniques, as well as the difficulty of each technique, already taxed their cognitive and attentional resources such that the added cognitive demands imposed by random practice did not help learning.

Consistent with the challenge point framework, our results also showed a learning advantage for individuals who practised using a self-directed schedule. This advantage over a matched control group has been previously shown for various aspects of practice[3–7] and can be ascribed to increased autonomy, which likely allows the participant to adapt the learning experience to his or her specific needs and may also result in increased motivation, more instances of deliberate practice and improved motor learning.[7] However, the advantage of a self-directed practice schedule over prescribed random and blocked practice is particularly interesting. Despite having chosen predominantly blocked schedules and changing to a random schedule later in the training phases, the self-directed group experienced superior learning. Furthermore, since the benefit of a self-directed practice schedule was significant only for the simplest suture technique, it is possible that increased task difficulty (and greater cognitive load) imposed by more difficult suture techniques reduced the advantage of a self-directed practice schedule for the horizontal and vertical mattress suture techniques. This emphasizes the complex interaction of task difficulty and training conditions that are required for optimal learning.[11] Nonetheless, the self-directed learning advantage that we observed is consistent with the literature that has been produced using basic laboratory tasks and now adds to the emerging work in the clinical skills domain, particularly for self-directed practice schedules as opposed to self-directed feedback or access to instructional materials.

Limitations

While this study is an important first step in understanding the role of practice schedules and instructional methods in learning surgical skills, we believe that more research is required to examine the impact of practice schedules for a variety of surgical tasks performed by surgeons with various skill levels. Furthermore, the generalizability of our results is limited to this particular skill, the population that was tested and the short time period over which we assessed learning effects. We are currently looking at similar processes in surgical residents to examine whether increased skill levels have any interaction with the already reported benefits to self-directed practice. Since our results were different across suturing techniques, we believe that in future studies and in practice, researchers and trainers should take care to examine component skills in a training program and so identify specific areas where trainees may require extra practice or instruction.

Conclusion

Many surgical skill centres are now offering 24-hour access to their facilities; however, instructors are often not available to provide expert direction during off-hours, leaving trainees to manage their own practice sessions. Our findings suggest that self-directed practice schedules within a curriculum may contribute to optimal learning of basic technical skills, such as simple suturing.

Competing interests: None declared.

Contributors: O. Safir, A. Dubrowski, D. Backstein and H. Carnahan designed the study. O. Safir acquired and analyzed the data, which was also analyzed by C.K. Williams. O. Safir, C.K. Williams, A. Dubrowski and H. Carnahan wrote the article. C.K. Williams, D. Backstein and H. Carnahan reviewed the article. All authors approved its publication.

Funding: This research was supported by a grant from the Royal College of Physicians and Surgeons of Canada.

References

1. Safir O, Dubrowski A, Hui Y, et al. Self-directed practice scheduling is equivalent to instructor guided practice when learning a complex surgical skill. *Procedia Soc Behav Sci* 2010;2:792-6.
2. Jowett N, LeBlanc V, Xeroulis G, et al. Surgical skill acquisition with self-directed practice using computer-based video training. *Am J Surg* 2007;193:237-42.
3. Chiviacowsky S, Wulf G. Self-controlled feedback: Does it enhance learning because performers get feedback when they need it? *Res Q Exerc Sport* 2002;73:408-15.
4. Chen DD, Hendrick JL, Lidor R. Enhancing self-controlled learning environments: the use of self-regulated feedback information. *J Hum Mov Stud* 2002;43:69-86.
5. Janelle CM, Barba DA, Frehlich SG, et al. Maximizing performance feedback effectiveness through videotape replay and a self-controlled learning environment. *Res Q Exerc Sport* 1997;68:269-79.
6. 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;43:507-15.

7. Wulf G, Raupach M, Pfeiffer F. Self-controlled observational practice enhances learning. *Res Q Exerc Sport* 2005;76:107-11.
8. Schmidt RA, Lee TD. *Motor learning and control*. 4th ed. Champaign (IL): Human Kinetics Publishers; 2005.
9. Shea JB, Morgan RL. Contextual interference effects on the acquisition, retention and transfer of a motor skill. *J Exp Psychol Learn Mem Cogn* 1979;5:179-87.
10. Lee TD, Magill RA. Can forgetting facilitate skill acquisition? In: Goodman D, Wilberg RB, Franks IM, editors. *Differing perspectives in motor learning, memory and control*. Amsterdam: Elsevier; 1985. p. 3-22.
11. Guadagnoli MA, Lee TD. Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *J Mot Behav* 2004;36:212-24.
12. Brady F. Contextual interference: a meta-analytic study. *Percept Mot Skills* 2004;99:116-26.
13. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. *Hum Mov Sci* 1990;9:241-89.
14. Dubrowski A, Backstein D, Abughaduma R, et al. The influence of practice schedules in the learning of a complex bone-plating surgical task. *Am J Surg* 2005;190:359-63.
15. Brydges R, Carnahan H, Backstein D, et al. Application of motor learning principles to complex surgical tasks: searching for the optimal practice schedule. *J Mot Behav* 2007;39:40-8.
16. Kurahashi A, Leming K, Carnahan H, et al. Effects of expertise, practice and contextual interference on adaptations to visuo-motor misalignment. *Stud Health Technol Inform* 2008;132:225-9.
17. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997;84:273-8.
18. Dubrowski A, Park J, Moulton C, et al. A comparison of single- and multiple-stage approaches to teaching laparoscopic suturing. *Am J Surg* 2007;193:269-73.
19. Downing SM. Reliability: on the reproducibility of assessment data. *Med Educ* 2004;38:1006-12.
20. Merbah S, Meulemans T. Learning a motor skill: effects of blocked versus random practice — a review. *Psychol Belg* 2011;51:15-48.