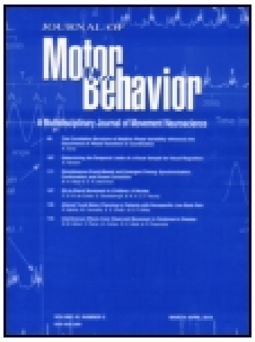

Routledge
Taylor & Francis Group

Journal of Motor Behavior

ISSN: 0022-2895 (Print) 1940-1027 (Online) Journal homepage: http://www.tandfonline.com/loi/vjmb20

Application of Motor Learning Principles to Complex Surgical Tasks: Searching for the Optimal Practice Schedule

Ryan Brydges , Heather Carnahan , David Backstein & Adam Dubrowski

To cite this article: Ryan Brydges , Heather Carnahan , David Backstein & Adam Dubrowski (2007) Application of Motor Learning Principles to Complex Surgical Tasks: Searching for the Optimal Practice Schedule, Journal of Motor Behavior, 39:1, 40-48, DOI: 10.3200/JMBR.39.1.40-48

To link to this article: http://dx.doi.org/10.3200/JMBR.39.1.40-48

Journal of Motor Behavior, 2007, Vol. 39, No. 1, 40–48

Application of Motor Learning Principles to Complex Surgical Tasks: Searching for the Optimal Practice Schedule

Ryan Brydges
Department of Surgery
Heather Carnahan
Department of Surgery
Department of Occupational Science
and Occupational Therapy
University of Toronto, Ontario, Canada

David Backstein
Adam Dubrowski
Department of Surgery
University of Toronto, Ontario, Canada

ABSTRACT. Practice of complex tasks can be scheduled in several ways: as whole-task practice or as practice of the individual skills composing the task in either a blocked or a random order. The authors used those 3 schedules to study 18 participants' learning of an orthopedic surgical task. They assessed learning by obtaining expert evaluation of performance and objective kinematic measures before, immediately after, and 1 week after practice (transfer test). During acquisition, the blocked group showed superior performance for simple skills but not for more complex skills. For the expert-based measures of performance, all groups improved from pretest to posttest and remained constant from posttest to transfer. Measures of the final product showed that the whole-practice group's outcomes were significantly better than those of the random group on transfer. All groups showed better efficiency of motions in the posttest than in the pretest. Those measures were also poorer on the transfer test than on the posttest. The present evidence does not support the contextual interference effect—hypothetically, because of the inherent cognitive effort effect associated with some of the component skills. The authors recommend that surgical tasks composed of several discrete skills be practiced as a whole. The results of this study demonstrate the importance of critically appraising basic theories in applied environments.

Key words: contextual interference, motor learning, practice schedule, skill acquisition

Theoretical advances in the principles governing motor skill acquisition are often used to guide interested educators and health professionals in applied settings to develop theory-based educational and rehabilitative goals. Conversely, careful implementation of the theoretical concepts in the applied world can serve as testing grounds that can lead to empirical validation of those concepts and to the generation of interesting theoretical questions.

One applied field in which theoretical motor learning principles can be tested is the training of technical surgical skills. The ability to adequately prepare new surgical trainees by using the traditional Halstedian apprenticeship model of surgical education—coined *see one, do one, teach one*—has been recently questioned. Demands on new surgical trainees are increasing because they are required to learn about more diseases and procedures in the same or fewer working hours. Learning opportunities are further limited by economic constraints in health care and by a greater respect for the rights of patients. It is becoming necessary for instructors to exploit the educational value of every learning experience in an effort to reduce the initial portion of the hypothetical learning curve for the technical aspects of the surgical craft (Hall, Ellis, & Hamdorf, 2003).

In an effort to standardize and optimize the learning experience, investigators have shown an increased interest in the role of surgical skills laboratories in the teaching of at least some of the most basic surgical skills (Aggarwal, Hance, & Darzi, 2004; Aggarwal, Moorthy, & Darzi, 2004). That environment provides an opportunity for trainees to practice technical tasks and skills repeatedly until they achieve proficiency. There is evidence that low-fidelity bench top models are as effective as high-fidelity models (Grober et al., 2004); the former models improve the cost-effectiveness of skills laboratories and maintain the learning objectives. However, the application of learning principles to the development of pedagogically sound and cost-effective practice schedules has been largely unexplored. That is, to date, efforts in surgical training have been focused on developing appropriate models for practice (Grober et al.) and methods

---

Correspondence address: Adam Dubrowski, University of Toronto, Surgical Skills Centre at Mount Sinai Hospital, 600 University Avenue, Level 2–Room 250, Toronto, Ontario, M5G 1X5, Canada. E-mail address: adam.dubrowski@utoronto.ca

of evaluating learning (Martin et al., 1997; Perkins, Starkes, Lee, & Hutchison, 2002). How those technical skills and tasks are best practiced, though, has not been well addressed within the literature.

In teaching complex surgical tasks composed of several technical skills, practice may be arranged either as *whole practice*—that is, the entire task is taught in its serial order—or as *part practice*—that is, the task is divided into its fundamental movement segments or skills (Dubrowski, Backstein, Abughaduma, Leidl, & Carnahan, 2005). Basic technical skills are thus the building blocks of tasks and can be objectively assessed in isolation (Bann, Khan, & Darzi, 2003; Carnahan, 1993; Datta et al., 2002). Pioneers in task analysis, Naylor and Briggs (1963) hypothesized that a task's complexity and its organization determine how it should be practiced. *Task complexity* has been defined as the number of movement segments (Magill, 2000), whereas *task organization* refers to the temporal relationship between the composite movement segments. According to that theoretical paradigm, optimal learning of a high-complexity, low-organization task will occur under part-practice conditions. Park, Wilde, and Shea (2004) described that method as one in which complex movement sequences (akin to technical surgical tasks) are decomposed into smaller, more manageable parts (akin to technical surgical skills) that one can later recombine to create a consolidated sequence. Another view is that improvement in the performance of sequential motor tasks can be regarded as the ability to perform the smaller subsequences together so that the transitions between the movement elements ultimately disappear (Hansen, Tremblay, & Elliott, 2005). In that view, whole practice would seem to be the preferred practice regime. In practical terms, however, part-practice, when conducted on surgically relevant bench models, is the most cost-effective choice because the trainees can share equipment, space, and the attention of the instructor.

When arranging a part-practice schedule, the instructor may teach individual technical skills in either a blocked or a random order. Comparison of those two practice schedules is extensive in motor learning research. Research in part-practice schedules has historically been focused on motor skills such as barrier knock-down (Shea & Morgan, 1979) and timing (Wulf & Lee, 1993) tasks. What differentiates blocked and random schedules is the amount of contextual interference (Battig, 1966): In the blocked schedule, contextual interference is low (Lee, Wishart, Cunningham, & Carnahan, 1997), whereas in the random schedule, contextual interference is high because that practice schedule requires that all skills be intermixed, creating a session without a definite pattern. The results of a majority of studies have demonstrated that although random practice is detrimental to performance during the initial practice sessions, it leads to better performance on retention and transfer tests than does blocked practice of the same skills (Magill & Hall, 1990). The contextual interference effect is not always present when studied within the realm of complex real-world tasks, however (Hebert, Landin, & Solmon, 1996). Wulf and Shea (2002) explained that the tasks commonly used to elicit the contextual interference effect are relatively simple, placing low demands on attention and memory and often requiring a small number of practice trials before performance achieves a steady level. Our unique aim in the present study was to address the issue of whether a complex surgical task can be learned efficiently by using the basic tenets of part–whole learning and contextual interference research.

In this study, we examined the differences in the acquisition of a bicortical bone-plating task imposed by different practice schedules. That surgical task can be defined as a *serial multisegmental task*. We subsequently divided the task into five fundamental technical skills. We isolated those skills for use in the part-practice conditions (Table 1). Although each skill is independent of the others, the order in which they must be conducted is crucial in the operating room.

Our purpose in the present study was to investigate part-versus-whole training in teaching the complex task of orthopedic bone plating. We used a transfer test to a more realistic model after a retention period to investigate the effect of practice conditions on the learning of the bone-plating task. We assumed that the transfer test would serve not only as a true indicator of actual learning (Schmidt & Bjork, 1992) but also as a good index of transfer of learning. Demonstrating transfer of learning from models with low fidelity (low difficulty) to human tissue (high difficulty) is highly sought in the surgical education domain. Based on earlier research (Dubrowski et al., 2005), our experimental hypothesis was that the blocked group would show superior performance during the acquisition phase, but that random and whole-practice schedules would lead to better transfer performance because those conditions contain a higher level of contextual interference.

Method

Participants

The University of Toronto Research Ethics Board approved this research, and we obtained informed consent from 18 postgraduate 1st-year surgical residents and 8 3rd-year undergraduate medical students (clerks). We randomly assigned participants to one of three groups; all orthopedic residents ($n = 6$) and clerks were equally distributed among each group.

Apparatus

We used anatomically correct foam cortical shell models (Pacific Research Laboratories, Sawbones, Vashon, WA) in the shape of a radius bone during the pre-, and posttests as well as during acquisition (Cristofolini & Viceconti, 2000; Szivek & Gealer, 1991; Szivek, Weng, & Karpman, 1990).

One of the five discrete skills required drilling actions (Table 1). Participants used a pneumatic drill (Hall Series 4, Model 5067, Zimmer Inc., Warsaw, IN) to perform that task. The other four skills composing the bone-plating task

TABLE 1. Description of Separate Skills Performed During the Practice Phase of the Experiment

Skill	Description
Reduction and application of plate	The (obliquely) fractured radius bone ends are opposed in an anatomical position and held with a bone clamp. Artificial soft tissues surrounded the bone. The bone plate is applied to that temporary fixation.
Drilling	This skill requires the appropriate use of a tissue guard and a power drill. The participant drilled five holes to accommodate the five screws required for the fixation. Each hole was drilled independently and with precise direction. The surgeon had to apply an appropriate amount of pressure to penetrate the bone while avoiding plunging.
Depth measurement	The performer uses a depth gauge to accurately measure each hole. The application of that device requires dexterity and the ability to feel the far cortex without actually seeing it (because of the soft tissue envelope).
Bone tapping	The participant performs this skill to produce threads for the screw. The technique requires reproduction of the identical angle used for the drilling procedure.
Insertion of screws	This skill requires the reproduction of the drilled hole's angle and the application of enough torque to enable the performer to insert all five screws tightly without stripping the threads.

Note. A panel of three orthopedic surgeons decomposed the entire bone-plating task into five functionally individual psychomotor skills.

were performed with the Zimmer ECT internal fracture fixation, small fragment plates and screws set (Zimmer Inc.). Participants used a cadaver specimen (whole arm) for the transfer test.

Procedure

Before the experimental session, all participants viewed a short instructional video in which an expert orthopedic surgeon slowly demonstrated the entire bone-plating task while explaining the component skills. The demonstration was error free. Then, all participants performed a pretest by completing the entire task, which consisted of proper bone reduction and plate application and the placement of five screws into the artificial bone. Next, they were randomly assigned to one of the three experimental groups. Each group performed all five skills involved in the bone-plating task (Table 1) in an assigned order. Whole training consisted of teaching all five component parts of the skill during each practice bout. Part training consisted of teaching each of the five component parts of the task separately. Furthermore, participants performed part training in either a blocked or random fashion. In blocked training, we taught one component five times before moving to the second component. In random training, we taught each component once, but the order of components was presented in a random fashion. Thus, the contextual interference (e.g., Shea & Morgan, 1979) varied between the two part-training methods.

We held the number of practice repetitions for each skill constant for all participants in all groups. One criticism of our method is that the number of trials was small. However, we intentionally selected that number to represent the most realistic practice schedule within the current junior residency educational program.

After the acquisition phase, participants performed a posttest on the same artificial radius. Following a 1-week retention period, participants returned to complete a transfer test on an artificially fractured cadaver radius (an expert surgeon cut the radius by using a pneumatic saw; Hall Series 4 oscillator). Thus, we assessed participants' performances three times: before acquisition (pretest); 5-min postacquisition (posttest); and, after a 1-week rest period, on cadaveric tissue (transfer test). During all three tests, it was not feasible to measure the five individual surgical skills separately. The measurements derived from those performances therefore represented the entire procedure.

Measurement

We measured technical performance during acquisition with the Imperial College (London) Surgical Assessment Device (ICSAD) motion analysis system, an objective method of quantifying the movement process of surgical skills (Aggarwal, Moorthy, et al., 2004). The ICSAD monitored hand motion characteristics by tracking the positions of magnetic markers placed on the dorsum of each participant's hands (Datta, Mackay, Mandalia, & Darzi, 2000). We used a commercially available motion-tracking system (Isotrak II, Polhemus, VT) to track the movement of the participants' instrumented hands in three-dimensional space coordinates. The sampling frequency of the system was 20 Hz. Using the position data, custom software (Imperial College, London) enabled us to derive two motion parameters online: (a) number of hand movements and (b) total time on task. Movement duration specifically has been considered the most sensitive measure of surgical competency on microscopic (Starkes, Payk, & Hodges, 1998) and laparoscopic (Perkins et al., 2002) suturing tasks.

We also captured the performances on all three tests on videotape for subsequent offline analysis by two expert orthopedic surgeons; the experts used three separate methods for evaluation of the tapes (Dath et al., 2004). The first was a modified version of a six-question, five-anchor-points global rating scale (GRS) developed to measure general operative performance (Martin et al., 1997). The second was a 15-item checklist of detailed, operation-specific procedures identified by a panel of three orthopedic surgeons as necessary to perform the operative task effectively. Those three surgeons were not otherwise involved in the study. The third expert-based evaluation method was a final product analysis, similar to the one developed by Szalay, MacRae, Regehr, and Reznick (2000) for end-to-side anastomosis.

As in the case of the duration of the acquisition phase, we selected all of those scoring systems because they are currently used during formative evaluation of surgical competency and would therefore likely enhance the applicability of our findings to the surgical education domain.

Statistical Analysis

Because of logistical constraints associated with expert rater availability, we applied only the computer-based assessment method during the acquisition phase of the experiment. We subjected number of movements (NM) and total time (TTime) to separate 3 (group: whole, random, blocked) × 5 (trial) mixed-design analyses of variance (ANOVAs) with repeated measures on the last variable. We analyzed the data for each of the five practiced skills in separate ANOVAs: (a) fracture reduction and application of plate, (b) drilling, (c) depth measurement, (d) bone tapping, and (e) insertion of screws.

We assessed the normality of the GRSs, the task-specific checklists, final product analysis, and ICSAD measures (NM and TTime) with a number of Shapiro–Wilk tests; the tests showed that all measures except the final product analysis were normally distributed ($p < .05$). On the basis of that result, we conducted appropriate statistical tests to examine whether the three groups were equivalent before any manipulations. For that purpose, we subjected the scores on the pretest to a one-way ANOVA and to a Kruskal–Wallis nonparametric test for final product analysis to ensure that all groups were equivalent before the acquisition phase. Where groups were found to differ at pretest, we calculated difference scores between the raw scores on the pre- and posttests and the pre- and transfer tests; we used the difference scores as an index of learning (Dubrowski et al., 2005).

The tests showed group equivalence for the checklist, $F(2, 23) = 0.72$, $p = .50$; GRS, $F(2, 23) = 1.26$, $p = .30$; NM, $F(2, 23) = 0.63$, $p = .54$; and TTime, $F(2, 23) = 0.43$, $p = .66$. However, there was a significant difference for the final product analysis, $\chi^2(4, N = 3) = 7.77$, $p < .05$. Post hoc analysis (Mann–Whitney *U* test) revealed that the random group ($M = 4.50$, $SE = 0.25$) had a significantly higher score than did both the whole group ($M = 3.33$, $SE = 0.38$; $U = 13.5$, $p < .05$) and the blocked group ($M = 3.63$, $SE = 0.23$; $U = 12.5$, $p < .05$) on pretest. Despite our best attempts at balancing the experimental groups on the basis of level of training, the pretest differences on that measure may be attributed to inadequate randomization or high variability in the final product analysis measure. We subsequently calculated difference scores for that dependent measure and used them for further group comparisons.

We analyzed all scores for the task-specific checklist, GRS, and the two kinematic variables (NM and TTime) in four separate repeated measures ANOVAs. Each ANOVA model consisted of a between-participants variable, 3 (group: whole, random, blocked), and a within-participants variable, 3 (test phase: pretest, posttest, and transfer test), with repeated measures on the last variable. For all ANOVAs, we further analyzed effects significant at $p < .05$ by using Tukey's honestly significant difference post hoc method for comparison of means. We analyzed the difference scores calculated for the final product analysis with the Kruskal–Wallis test to investigate main effects, and we used the Mann-Whitney *U* test to perform post hoc analysis.

Results

Acquisition Phase

Computer-Based Assessment

Table 2 illustrates the group and trial main effects for the two computer-based assessment variables measured during acquisition. No group main effects were found for the reduction and application of plate, depth measurement, bone-tapping, or insertion of screws skills. For the drilling skill, post hoc analysis on the group main effect indicated that the blocked group performed fewer movements (i.e., more efficient execution) and required less time to complete the skill than did the other groups, and that the whole and random groups did not differ. The depth measurement and bone-tapping skills did not show any significant trial main effects for either of the two measures. For the insertion of screw skill, there was no trial main effect for the NM variable. However, the time to complete the skill did decrease between Trial 1 and Trial 5 ($p < .05$). There was similarly a statistically significant difference for both computer-based measures (indicating improvement) between Trial 1 and Trial 5 for both the reduction and application of plate and the drilling skills. No Group × Trial interactions were found during the acquisition phase.

Testing Phase

Computer-Based Measures

The analyses of the number of movements revealed a main effect only for test, $F(1, 23) = 11.71$, $p < .01$ (Figure 1A). All groups improved their performance from pretest to posttest. That measure was also sensitive to changes in performance during the 1-week rest period. There was an increase in the number of movements from posttest to transfer test (Figure 1A).

Analyses of TTime similarly revealed a main effect only for test, $F(1, 23) = 41.40$, $p < .001$ (see Figure 1B); all groups improved their performance from pretest to posttest. That measure was also sensitive to changes in performance during the 1-week rest period. There was an increase in

TABLE 2. Acquisition Findings for Each Individual Psychomotor Skill

	No. of movements			Time taken (s)		
Main effect	*F*	*df*	*p*	*F*	*df*	*p*
			Reduction and application of plate			
Group	1.40	2, 23	.27	0.94	2, 23	.40
Trial	14.74	1, 23	< .001	15.93	1, 12	< .001
			Drilling			
Group	9.62	2, 23	< .001	3.51	2, 23	< .05
Trial	5.65	1, 23	< .05	13.51	1, 23	< .001
			Depth measurement			
Group	2.51	2, 23	.10	3.16	2, 23	.06
Trial	3.38	1, 23	.08	1.60	1, 23	.22
			Bone tapping			
Group	0.11	2, 23	.89	0.35	2, 23	.71
Trial	0.56	1, 23	.46	0.06	1, 23	.82
			Insertion of screws			
Group	1.98	2, 23	.16	0.16	2, 23	.85
Trial	0.88	1, 23	.36	5.58	1, 23	< .05

TTime from posttest to transfer test (Figure 1B). It is not clear whether those increases in number of movements and total time on task were a result of skill degradation or whether they were a function of the more realistic model used for the transfer test.

Expert-Based Measures

Analysis of the GRS revealed a main effect for test, $F(1, 23) = 33.31$, $p < .001$ (see Figure 1C); all three groups improved their performance from the pretest to both the posttest and the transfer test, with no statistically significant differences between the post- and transfer tests. The analyses of the practice-specific checklists similarly revealed a main effect for test, $F(1, 23) = 33.55$, $p < .001$ (Figure 1D). All groups showed improvements in scores from the pretest to both the posttest and the transfer test. There were no interactions between the experimental groups and test phase for either the checklist or the GRS.

Because there was a significant difference between the random group and the two other groups on pretest for the final product analysis, we used individual difference scores for that measure. The change in the final product analysis from pretest to posttest did not yield statistically significant differences across the three experimental groups, $\chi^2(2, N = 3) = 4.39$, $p = .11$. However, the change from pretest to transfer test was significantly different, $\chi^2(2, N = 3) = 5.91$, $p < .05$. Here, the whole group demonstrated greater improvements on the final product than did the random group, and the improvements demonstrated by the blocked group were equivalent to those of both the whole and the random groups (Figure 1E).

Discussion

In this study, we focused on investigating the most appropriate practice schedule for an orthopedic surgical task. That issue is of utmost importance for individuals who teach surgical technical skills, who are always searching for ways to enhance the teaching capabilities of staff surgeons and the learning opportunities of residents in laboratory-based teaching venues to improve patient care delivery and safety (Dubrowski et al., 2005; Hamstra & Dubrowski, 2005). Our aim in the current investigation was to determine the most beneficial learning paradigm for novice surgeons practicing a complex bone-plating task, with the ultimate goal of optimizing the learning experience in the laboratory setting. We used two well-known theoretical principles from the motor learning literature to achieve that goal, namely, the part–whole practice paradigm and contextual interference.

Acquisition Phase

Our results in the acquisition phase showed contextual interference for only a subset of skills; the relative complexity of each of the five skills composing the bone-plating task can explain that finding. Drilling and depth measure-

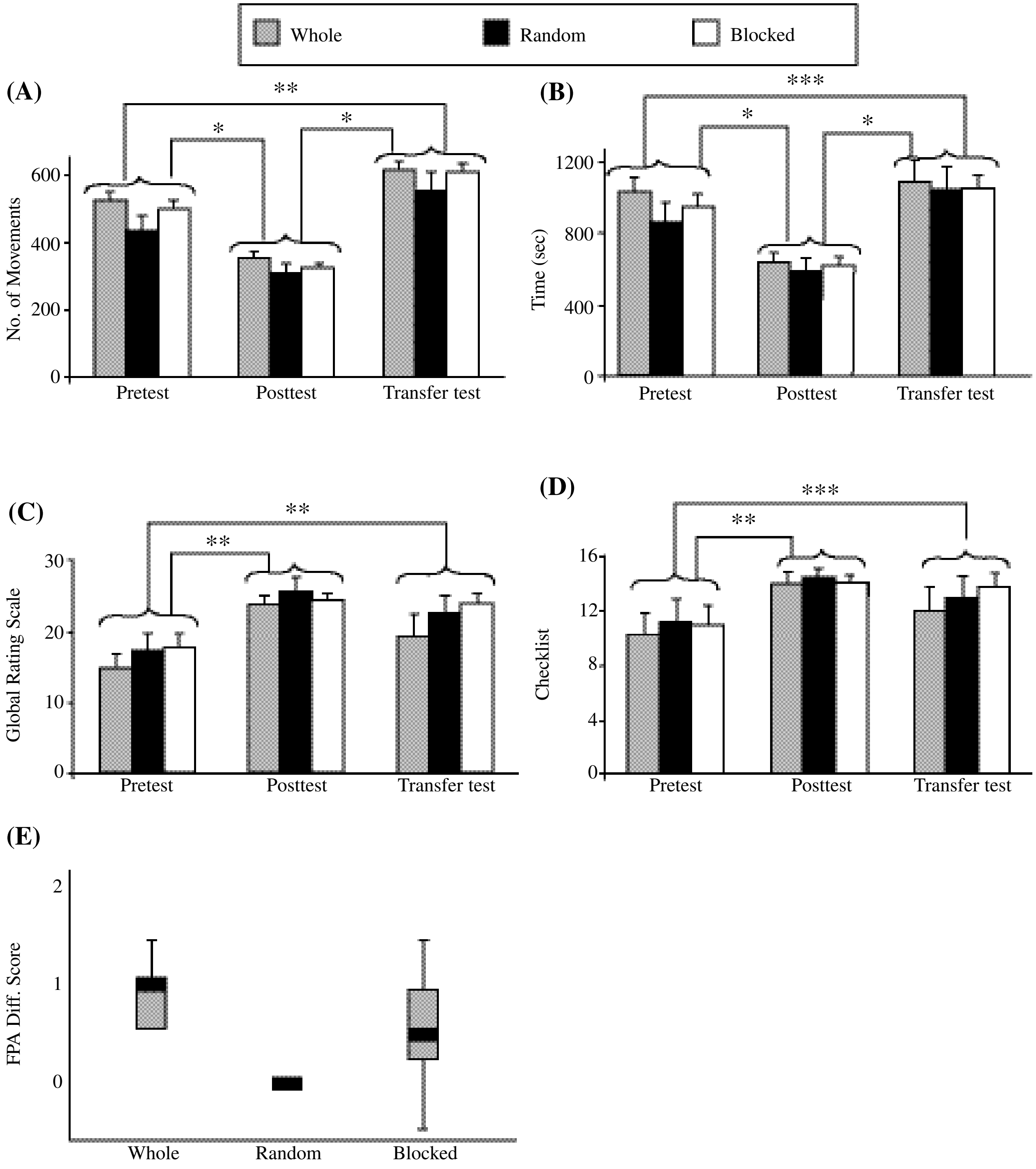

FIGURE 1. Significant effects for all computer- and expert-based measures. The computer-based measures, (**A**) number of movements and (**B**) time on task, are both charted as a function of test. The expert-based measures, (**C**) global rating scale and (**D**) checklist score, are charted across test. The difference score between the pretest and the transfer test, which was calculated for the (**E**) final product analysis (FPA Diff. Score), is charted across group. $^*p < .001$. $^{**}p < .01$. $^{***}p < .05$.

ment showed specific improvements in performance as a function of practice for the blocked group over those of the random and whole groups, whereas blocked practice did not differentially improve the reduction and application of the plate, bone-tapping, and insertion of screws skills. One plausible explanation for the differential effect is the inherent cognitive effort involved in producing a particular skill. One theoretical explanation of the contextual interference effect suggests that the high cognitive effort associated with practicing under random conditions promotes later retention and transfer of the practiced task (Lee & Simon, 2004; Schmidt & Lee, 2005). However, Albaret and Thon (1998) put forward the proposition that as the complexity of the practiced skill increases, participants are forced to use intensive cognitive processing and recurrent calls on long-term memory, regardless of the practice conditions. For example, the information in a participant's working memory could undergo decay because of the processing of feedback received while completing each complex submovement, forcing him or her to reconstruct the action plan on the subsequent trial even if the skill does not change (Albaret & Thon). Thus, the intraskill interference created by the inherent complexity of the skill may obscure the contextual interference effect because the increase in cognitive load could possibly override the benefits of the interskill variation produced by a random practice schedule. Moreover, the combination of random practice and effectively complex submovements could overload the system, reducing the amount learned within that schedule (Wulf & Shea, 2002).

Therefore, we first suggest that the component skills of the bone-plating task may be categorized as simple or complex. For instance, the drilling and depth measurement skills may be categorized as simple because the movements require a single tool, only one to two degrees of freedom, and minimal coordination between the joints of the upper limb. Those tasks are inherently similar to the simple motor tasks (e.g., key press sequences) commonly employed in prominent contextual interference and part–whole practice experiments (Hansen et al., 2005; Shea & Morgan, 1979). On the other hand, the skills of reduction and application of plate, bone-tapping, and insertion of screws are complex because they require the use of multiple tools, temporal coordination of upper limb joints, and movements through several degrees of freedom. Our post hoc categorization supports Albaret and Thon's (1998) contention that skill complexity interferes (interacts or masks) with the contextual interference effect. One limitation of the present design is that we decomposed the task of bone-plating to five skills (movement segments). We based our selection of that decomposition of the task on consultations with practicing orthopedic surgeons responsible for teaching the task in the operating room to novice trainees and on the physical arrangements used in laboratory-based training, which are dictated mainly by equipment and space limitations. Therefore, the task decomposition was functional. We recognize that, with further decomposition of the complex skills into more refined and simpler skills, the results of the acquisition phase could possibly be different and more in line with the contextual interference effect.

Testing Phase

The present results suggest that the combined effect of the intraskill difficulties within the three complex skills led to an increase in the cognitive load experienced by participants in all three experimental groups. Our results support the notion that an increase in contextual interference facilitated retention in all groups, as evaluated by the computer-based assessment and by two of the expert-based assessments. That finding differs from the results reported by Dubrowski et al. (2005), who conducted a similar study with medical students. In their study, the whole and random groups performed better than did the blocked group on a retention test (i.e., the same model was used across tests). Indeed, nonsignificant trends in the expert-based transfer test data suggested similar results, with the whole and random groups scoring better than the blocked group on the GRS and checklist measures (Table 3).

An alternative explanation of the transfer test findings makes use of tenets from the optimal framework proposed by Guadagnoli and Lee (2004). In their challenge-point framework, they proposed that every motor task can possess two types of difficulties: nominal and functional. The nominal task difficulty is the complexity of the task under optimal conditions, and the functional difficulty of the task is the difficulty modulated by external factors. For example, performing an orthopedic bone-plating task entails a certain amount of nominal difficulty. Performing the same task on an inanimate model in a laboratory environment results in a functional difficulty lower than the nominal difficulty because of the less stressful environment (Hauge, Wanzek, & Godellas, 2001; Lingard, Reznick, Espin, Regehr, & DeVito, 2002). Performing the task on a cadaveric model may result in a higher functional difficulty, and performing the task in the operating room will further increase the level of functional difficulty. One of the premises of Guadagnoli and Lee's framework is that to evaluate learning objectively, one must adjust the functional difficulty of the task to the trainees' current performance level. In the present study, we trained the participants during the acquisition phase by using a low-fidelity model of a fractured radius bone. Guadagnoli and Lee would describe that task as having low-to-moderate functional difficulty because the model would present a relatively small challenge to the tested population. The level of functional task difficulty was significantly increased, however, on the transfer test because the use of the cadaveric arm introduced a number of new challenges, including a new

TABLE 3. Results of Transfer Test Group Comparisons

	Result		
Dependent variable	*F*	*df*	*p*
Number of movements	0.31	2, 23	.74
Time taken	0.31	2, 23	.95
Global rating scale	1.94	2, 23	.17
Checklist	2.39	2, 23	.11

Note. The final product analysis (difference score) was $\chi^2(2, N = 3) = 5.91$, $p < .05$.

sense of realism, wet surfaces, visual obstruction of the operating field by tissue surrounding the bone, denser bone tissue, and variability in the size and shape of the bone. According to the challenge-point framework, the amount of potentially available information in a task increases exponentially for novice performers with small increases in the functional difficulty. Thus, we propose that because our acquisition phase did not advance the capabilities of our performers far enough beyond the level of novice, their performance on the transfer test would be quite poor because of the high internal complexity of the task. That result was confirmed—all participants performed at a degraded level on the transfer test regardless of the practice schedule they followed.

Measurement Systems

There was no conformity between the process-oriented (e.g., computer-based measures, GRS, and task-specific checklist) and the outcome-oriented (e.g., final product analysis) dependent measures. Although the process-oriented measures showed that the learning of all groups was indistinguishable when assessed on the transfer test, it remains significant that the outcome-oriented measure was sensitive in detecting group differences following the retention period. That is, the improvements in the final product analyses from pretest, through posttest, to transfer test showed that the whole-practice group significantly improved their outcome in comparison with that of the random group. That outcome measure is, however, potentially limited by a ceiling effect because the scale of the final product analysis ranged from 0 to 5. The random group obtained an improvement score of only 0.11 because the group averaged similar values on both the pre- and transfer tests. On the basis of that observation, we feel that in future studies in which final product analysis is used as a dependent variable, researchers should consider a larger and more sensitive scale. Also, the significant group difference at outcome would have found greater support in this study had we used another computer-based outcome-oriented measure to evaluate the plated bone. To account for that limitation, in future studies investigators should find concomitant computer-based outcome-oriented measures for the applied task being investigated.

Practical Implications

Despite the aforementioned limitations, our results suggest that whole practice is the most beneficial schedule for the complex bone-plating task. When placed on the contextual interference continuum, that schedule would likely cause medium levels of contextual interference, which appear to be sufficient for improving performance.

On the basis of our results, it appears that bone-plating is a task of high organization and high complexity (Naylor & Briggs, 1963). That finding suggests that the composite skills must be practiced in whole order—that is, under whole practice conditions—probably because the transition between skills creates a significant change in the kinematic characteristics of each component. Wenderoth, Puttemans, Vangheluwe, and Swinnen (2003), who studied the effectiveness of part-versus-whole training on learning a bimanual task, provided support for that view. They concluded that whole-practice conditions enhance the learning of tasks involving a high degree of interlimb coordination, such as athletic or musical tasks. Moreover, Hansen et al. (2005) discussed a phenomenon termed the *one-target advantage*, showing that when individuals string together subcomponents of a task to perform a complete movement series, the kinematic characteristics of each skill often change because of planning or initiation of adjacent movement elements. Both of those examples suggest that the order of learning of submovements (here, termed *technical skills*) of a complex motor task is important to the overall performance of the task. Evidence is mounting in support of practice schedules that require the entire movement sequence to be performed during acquisition, especially for real-world tasks that are inherently complex. Furthermore, the results indicate the applicability of Guadagnoli and Lee's (2004) challenge-point framework in the surgical domain. The results of the present study have specifically shown the importance of matching the practice schedule to the performer, environment, and task. In general, surgical educators can exploit this framework to determine how practice should be set up for efficient learning by the novice trainee.

Theoretical Contributions

From a theoretical perspective, the debate regarding the transferability of principles derived from studies using simple skills to complex skill learning is also of significant interest. Wulf and Shea (2002) speculated that one feature of motor learning that is difficult for experimenters to simulate by using simple laboratory skills is that the learner's attention may be directed to several task elements. For example, the skills practiced in this study consisted of large compound movements whose subcomponents differed in many characteristics and often involved a total movement time of 10–30 s during just one trial. Most motor-learning principles have been developed on the basis of less complex movements with relatively fewer degrees of freedom and vastly shorter movement times (on the millisecond scale). We believe that the use of existing theories and principles in complex environments plays a significant role in knowledge translation from theoretical to applied fields, and vice versa. In particular, principles from the challenge-point framework (Guadagnoli & Lee, 2004) have been useful in understanding the changes in motor behavior caused by manipulation of practice schedules in the applied surgical setting. There appears to be a necessary requirement for researchers interested in motor learning to apply long-standing theories in more naturalistic settings while preserving the scientific methods of the past (Wulf & Shea).

REFERENCES

Aggarwal, R., Hance, J., & Darzi, A. (2004). Surgical education and training in the new millennium. *Surgical Endoscopy, 18,* 1409–1410.

Aggarwal, R., Moorthy, K., & Darzi, A. (2004). Laparoscopic skills training and assessment. *British Journal of Surgery, 91,* 1549–1558.

Albaret, J. M., & Thon, B. (1998). Differential effects of task complexity on contextual interference in a drawing task. *Acta Psychologica, 100,* 9–24.

Bann, S. D., Khan, M. S., & Darzi, A. (2003). Measurement of surgical dexterity using motion analysis of simple bench skills. *World Journal of Surgery, 27,* 390–394.

Battig, W. F. (1966). Facilitation and interference. In E. A. Bilodeau (Ed.), *Acquisition of skill* (pp. 134–146). New York: Academic Press.

Carnahan, H. (1993), The role of three-dimensional movement analysis in the assessment of motor expertise. In J. Starkes & F. Allard (Eds.), *Cognitive issues in motor expertise* (pp. 35–53). Amsterdam: Elsevier.

Cristofolini, L., & Viceconti, M. (2000). Mechanical validation of whole bone composite tibia models. *Journal of Biomechanics, 33,* 279–288.

Dath, D., Regehr, G., Birch, D., Schlachta, C., Poulin, E., Mamazza, J., et al. (2004). Toward reliable operative assessment: The reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surgical Endoscopy, 18,* 1800–1804.

Datta, V., Mackay, S., Mandalia, M., & Darzi, A. (2000). The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of the American College of Surgeons, 193,* 479–485.

Datta, V., Mandalia, M., Mackay, S., Chang, A., Cheshir, N., & Darzi, A. (2002). Relationship between skill and outcome in the laboratory-based model. *Surgery, 131,* 318–323.

Dubrowski, A., Backstein, D., Abughaduma, R., Leidl, D., & Carnahan, H. (2005). The influence of practice schedules in the learning of a complex bone-plating surgical task. *American Journal of Surgery, 190,* 359–363.

Grober, E. D., Hamstra, S. J., Wanzel, K. R., Reznick, R. K., Matsumoto, E. D., Sidhu, R. S., et al. (2004). The educational impact of bench model fidelity on the acquisition of technical skill: The use of clinically relevant outcome measures. *Annals of Surgery, 240,* 374–381.

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects for various practice conditions in motor learning. *Journal of Motor Behavior, 36,* 212–224.

Hall, J. C., Ellis, C., & Hamdorf, J. (2003). Surgeons and cognitive processes. *British Journal of Surgery, 90,* 10–16.

Hamstra, S. J., & Dubrowski, A. (2005). Effective training and assessment of surgical skills, and the correlates of performance. *Surgical Innovation, 12,* 71–77.

Hansen, S., Tremblay, L., & Elliott, D. (2005). Part and whole practice: Chunking and online control in the acquisition of a serial motor task. *Research Quarterly for Exercise and Sport, 76,* 60–66.

Hauge, L. S., Wanzek, J. A., & Godellas, C. (2001). The reliability of an instrument for identifying and quantifying surgeons' teaching in the operating room. *American Journal of Surgery, 181,* 333–337.

Hebert, E. P., Landin, D., & Solmon, M. A. (1996). Practice schedule effects on the performance and learning of low- and high-skilled students: An applied study. *Research Quarterly for Exercise and Sport, 67,* 52–58.

Lee, T. D., & Simon, D. A. (2004). Contextual interference. In A. M. Williams (Ed.), *Skill acquisition in sport: Research, theory and practice* (pp. 29–44). London: Routledge.

Lee, T. D., Wishart, L. R., Cunningham, S., & Carnahan, H. (1997). Modeled timing information during random practice eliminates the contextual interference effect. *Research Quarterly for Exercise and Sport, 68,* 100–105.

Lingard, L., Reznick, R., Espin, S., Regehr, G., & DeVito, I. (2002). Team communications in the operating room: Talk patterns, sites of tension, and implications for novices. *Academic Medicine, 77,* 232–237.

Magill, R. A. (2000). *Motor learning: Concepts and applications*. New York: McGraw-Hill.

Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effects in motor skill acquisition. *Human Movement Science, 9,* 241–289.

Martin, J. A., Regehr, G., Reznick, R., MacRae, H., Murnaghan, J., Hutchison, C., et al. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery, 84,* 273–278.

Naylor, J. C., & Briggs, G. E. (1963). Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *Journal of Experimental Psychology, 65,* 217–244.

Park, J. H., Wilde, H., & Shea, C. H. (2004). Part–whole practice of movement sequences. *Journal of Motor Behavior, 36,* 51–61.

Perkins, N., Starkes, J. L., Lee, T. D., & Hutchison, C. (2002). Learning to use minimal access surgical instruments and 2-dimensional remote visual feedback: How difficult is the task for novices? *Advances in Health Sciences Education: Theory and Practice, 7,* 117–131.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualization of practice: Common principles in three research paradigms suggest important new concepts for practice. *Psychological Science, 3,* 207–217.

Schmidt, R. A., & Lee, T. D. (2005). *Motor control and learning: A behavioral emphasis* (4th ed.). Champaign, IL: Human Kinetics.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology, 5,* 179–187.

Starkes, J. L., Payk, I., & Hodges, N. J. (1998). Developing a standardized test for the assessment of suturing skill in novice microsurgeons. *Microsurgery, 18,* 19–22.

Szalay, D., MacRae, H., Regehr, G., & Reznick, R. (2000). Using operative outcome to assess technical skill. *American Journal of Surgery, 180,* 234–237.

Szivek, J. A., & Gealer, R. L. (1991). Comparison of the deformation response of synthetic and cadaveric femora during simulated, one-legged stance. *Journal of Applied Biomaterials, 2,* 277–280.

Szivek, J. A., Weng, M., & Karpman, R. (1990). Variability in the torsional and bending response of a commercially available composite femur. *Journal of Applied Biomaterials, 1,* 183–186.

Wenderoth, N., Puttemans, V., Vangheluwe, S., & Swinnen, S. P. (2003). Bimanual training reduces spatial interference. *Journal of Motor Behavior, 35,* 296–308.

Wulf, G., & Lee, T. D. (1993). Contextual interference in movements of the same class: Differential effects on program and parameter learning. *Journal of Motor Behavior, 25,* 254–263.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin and Review, 9,* 185–211.

Submitted December 12, 2005
Revised June 13, 2006
Second revision July 18, 2006