The Effects of Spacing, Naps, and Fatigue on the Acquisition and Retention of Laparoscopic Skills ☆, ☆☆ ☆

Edward N. Spruit, MSc,[*,†,‡] *Guido P.H. Band, PhD,*[*,†] *Kristiaan B. van der Heijden, PhD,*[†,§] *and Jaap F. Hamming, PhD*[‡]

[*]Cognitive Psychology, Institute of Psychology, Leiden University, Leiden, The Netherlands; [†]Leiden Institute for Brain and Cognition (LIBC), Leiden, The Netherlands; [‡]Department of Surgery, Leiden University Medical Center (LUMC), Leiden, The Netherlands; and [§]Department of Clinical Child and Adolescent Studies, Institute of Education and Child Studies, Leiden University, Leiden, The Netherlands

BACKGROUND: Earlier research has shown that laparoscopic skills are trained more efficiently on a spaced schedule compared to a massed schedule. The aim of the study was to estimate to what extent the spacing interval, naps, and fatigue influenced the effectiveness of spacing laparoscopy training.

METHODS: Overall 4 groups of trainees (aged 17-41 y; 72% female; $N_{massed} = 40$; $N_{break} = 35$; $N_{break-nap} = 37$; $N_{spaced} = 37$) without prior experience were trained in 3 laparoscopic tasks using a physical box trainer with different scheduling interventions. The first (massed) group received three 100-minute training sessions consecutively on a single day. The second (break) group received the sessions interrupted with two 45-minute breaks. The third (break-nap) group had the same schedule as the second group, but had two 35-minute powernap intervals during the breaks. The fourth (spaced) group had the 3 sessions on 3 consecutive days. A retention session was organized approximately 3 months after training.

RESULTS: The results showed an overall pattern of superior performance at the end of training and at retention for the spaced group, followed by the break-nap, break, and massed group, respectively. The spaced and break-nap group significantly outperformed the break and massed group, with effect sizes ranging from 0.20 to 0.37.

CONCLUSIONS: Spacing laparoscopic training over 3 consecutive days or weeks is superior to massed training, even if the massed training contains breaks. Breaks with sleep opportunity (i.e., lying, inactive, and muted sensory input) enhance performance over training with regular breaks and traditional massed training. (J Surg Ed ∎:∎∎∎-∎∎∎. © 2016 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: laparoscopy, simulation training, spacing, powernap, consolidation, motor skills

ACGME COMPENTENCIES: Practice-based Learning & Improvement

INTRODUCTION

Acquiring laparoscopic motor skills requires an extensive amount of practice. Earlier research suggested that practice time allocated for surgical training can be used most efficiently when scheduled across multiple smaller time intervals,[1,2] preferably with several nontraining days in between training sessions. In an earlier study,[3] we established that laparoscopic skills are better acquired and retained when learned on a spaced schedule as compared to a massed schedule. In the study, 2 groups of participants learned 4 basic and 1 advanced task (intracorporeal suturing) on a physical box trainer. The first group received training on a massed 1-day schedule, whereas the second group received the same amount of training divided over 3 consecutive weekly sessions. Performance at the end of training, at a 2-week retention session and a 1-year retention session, was superior for the spaced group. These spacing effects were most pronounced for the advanced task.

Spacing training provides multiple benefits. Having multiple shorter training sessions reduces the likelihood of trainees becoming overly fatigued[4] or bored that may occur

when practicing laparoscopic tasks for an extended amount of time. Also, trainers and trainees get a more accurate reflection of their actual skill level during spaced training.[5] As the training uses multiple sessions, trainees will discover that they do not perform as smoothly at the start of the second training session as they did at the end of the first training session. This gives them a more accurate appraisal of their own skill level. The fresh start at the beginning of each new training session requires trainees to invest more effort to get their performance back up to the level they left at the end of the previous session. This allows for more elaborate and frequent activations of the neural pathways in the brain associated with learning the specific motor skill.

Theorists have suggested that consolidation of memory is adaptive,[6] and that learning a skill is favored on a spaced training schedule because frequent short practice sessions on a training task give the brain the indication that it will encounter the same task more often in the future. This triggers more enduring memory encoding to accommodate for repetitive encounters with the task. Spacing provides time intervals between training sessions during which memory consolidation[7] can be strengthened. This in turn supports longer retention. Consolidation occurs in the brain when a person is disengaged from the trained activity and is enhanced during sleep,[8] when perceptual input to the brain is reduced to a large extent. In designing training, one ought to be cognizant of the fact that most learning takes place during off-line periods, not during practice. During training, practice provides the learning input that will be consolidated at a later period of time posttraining. If too many training tasks are practiced right after another, the learning input created by the first training experience will be replaced and thus impaired by later training experiences.[7] A more adequate strategy is to provide trainees rest intervals so the associated brain regions can have the opportunity to process the learning input before proceeding with more training. Research has shown that consolidation is enhanced by overnight sleep[9] as well as power naps.[10]

From our earlier results,[3] we were unable to distinguish to what extent each of these factors facilitates acquisition and retention of laparoscopic motor skills acquired on a physical box trainer. In the current study, we use a more elaborate design to dissect the individual effects that might contribute to spacing training.

The goal of this study was 2-fold: (1) to compare the effectiveness of an intervention with a smaller spacing interval to a bigger one (1 d versus 1 wk) and (2) to determine the influence of factors such as naps, mental fatigue, and simple breaks on the skill acquisition and retention of laparoscopy training. With the current study, we aimed to attain valid estimates of the extent to which each of these factors influences the effectiveness of spacing training.

First, we hypothesize that groups that have a larger time interval between training sessions will have superior performance at the end of training and at retention and experience lower levels of fatigue. Second, we hypothesize that sleep opportunities in between training sessions will lead to superior performance at the end of training and at retention and result in lower levels of fatigue.

METHODS

Participants

A total of 149 university students (108 female) without any prior experience in laparoscopy training were enrolled in the study. Age ranged from 17 to 41 years (mean = 21), and 128 participants were right-handed. All subjects filled out informed consent forms and were granted a training certificate as a reward for participating in the study.

Apparatus

Participants trained 3 laparoscopic tasks on a physical box trainer. All tasks have previously established construct validity.[11] The tasks train perceptual and motor skills such as depth perception, adapting to the fulcrum effect, and instrument handling. In the first task, participants had to coordinate a pipe cleaner through a set of 4 rings. This task is used to improve a trainee's bi-manual dexterity. In the second task, participants had to pick up small beads from a bucket and drop them on pins on a pegboard making the shape of a simple figure. This task requires caution and very careful handling of the pins with the instruments. In the final task, participants learn intra-corporeal suturing. The suturing sequence required correct insertion of the needle and 3 knots. The first knot required 2 throws, whereas the second and the third knot required 1 throw. Participants were instructed to start with the needle in their right instrument for the first and the third knot and to start with the needle in their left instrument for the second knot. This was done to ensure flexibility of the skill for both hands (see online Video Appendix for all the laparoscopic training tasks). During training, participants also learned suturing on an open model in order to prepare them before practicing in the laparoscopic box trainer.

Video footage of the performance of participants on the box-trainers was converted to .mpg files for each task at each moment of measurement using a video splitter, grabster, (Terratec Grabster AV 400 MX) and VLC Media Player for Windows.

Participants filled out self-report questionnaires covering demographics (sex, age, etc.), prior sport, music and gaming experience (0 = no experience, 1 = I used to play, 2 = yearly, 3 = monthly, 4 = weekly, and 5 = daily), personality,[12] and mental fatigue (rating scale mental effort [RSME][13]).

Training Programs

Training was divided into 3 sessions. The duration of each session was 100 minutes, divided into 50 minutes of practice, 15 minutes of instructions, and 35 minutes of measurement (laparoscopic tasks, questionnaires). All 3 laparoscopic tasks were practiced and measured during each of the 3 sessions, but the duration of practice for the pipe cleaner and beads task was reduced for the later sessions, while practice time on intracorporeal suturing was increased. Total practice time for the pipe cleaner, beads, and intracorporeal suturing task was 27.5, 32.5, and 60 minutes, respectively. During the first session, measurement took place before practice to establish a baseline. At the baseline measurement for intracorporeal suturing, only the insertion of the needle in the model is performed, as the knot tying part of the task is too difficult for a novice without any prior training. For the second and the third session, measurement took place after practice at the end of the session. RSME questionnaires were filled out right before measuring the laparoscopic tasks on each session. A retention session was scheduled approximately 3 months after training. All 3 laparoscopic tasks were measured during retention without any prior practice.

Participants received standardized instructions by the trainer and instructional videos (see Online Video Appendix) and no feedback in order to minimize confounding effects on the learning curve of the trainees. If trainees asked for feedback, they were reminded to pay close attention next time the instructional video would be shown, as all the required information to learn is present in the video.

There were 4 groups, each with a different time schedule. The spaced group received the training sessions on 3 consecutive days. The massed group had all sessions consecutively on a single day without any breaks. The break and break-nap group received the sessions on 1 day with two 45-minute breaks in between, with the break-nap group having a 35-minute powernap opportunity during both breaks. In the break-nap group, participants had a powernap opportunity on inflatable mattresses in a dark room while wearing earplugs and a sleeping mask to minimize sensory input. After each powernap, a self-report sleep questionnaire was filled out. All participants in the break and break-nap group wore activity trackers (Flex, Fitbit[14]) on their wrist to measure their activity levels during training and during the breaks. The instrument uses 3-dimensional motion sensing technology, records data in 1-minute epochs, and has shown to be an acceptable instrument for sleep/wake monitoring in normative populations.[14] Participants were assigned quasi-randomly to each group. Each group had randomly assigned timeslots allocated to them (unknown to the participants), and participants signed up for timeslots according to their availability. All groups trained their laparoscopic skills for an equal amount of time.

Performance Outcome Measures

The video files of the participants were assessed by the first author for completion times of the task. If participants were unable to complete the task within a set amount of measurement time (maximum of 10 min during training and 15 min during retention), a score of 601 or 901 (seconds) was assigned to avoid selective drop-out from our sample based on poor performance. This score would automatically be assigned as the highest rank in the nonparametric tests. Thus, there were 3 outcome measures (pipe-cleaner, beads, and suturing task) at each moment of measurement (at baseline, end of session 2, at the end of training, and retention).

A total score was computed as a fourth measure using z-scores from the laparoscopic tasks (we chose this method to make sure all 3 tasks contribute equally to the total score, as the laparoscopic tasks have differing mean completion times).

Statistical Analysis

Data were checked for normality, and statistical tests were chosen accordingly (one-way analysis of variance, independent samples, and paired samples t-tests in the case of normal data and Kruskal-Wallis, Mann-Whitney, and Wilcoxon Signed-rank tests in the case of nonnormal data) using the statistical software SPSS 23.0. A significance level of 0.05 was used. We tested whether groups were comparable at baseline in terms of age, sex, hand preference, academic year, food, caffeine, alcohol intake, sleep before training, mental fatigue, musical, gaming, sports activity, and personality factors. For our main analysis, we tested all 4 laparoscopic scores (pipe cleaner, beads, suturing, and total scores) on all 4 moments of measurement (baseline, at the end of session 2, at the end of training, and at the retention session) with 16 Kruskal-Wallis tests.

RESULTS

A total of 149 participants ($N_{massed} = 40$; $N_{break} = 35$; $N_{break-nap} = 37$; $N_{spaced} = 37$) completed the training and 134 participants ($N_{massed} = 36$; $N_{break} = 27$; $N_{break-nap} = 35$; $N_{spaced} = 36$) returned to the laboratory for the follow-up retention session.

Participant Characteristics

Chi-square tests showed no significant differences between all the groups in terms of sex, hand preference, and food intake. Caffeine intake was lower in the break-nap group (35.1%) compared to the other groups (massed: 62.5%; break: 51.4%; spaced: 70.2%; p = 0.022). Alcohol intake (on the night before training) differed between groups (massed: 62.5%; break: 37.1%; break-nap: 35.1%; spaced: 13.5%; p < 0.001). To check for any effects of caffeine or alcohol intake on performance on the laparoscopic tasks, 8 Mann-Whitney tests were performed (2 within each of the 4 groups). None of the tests were significant, indicating

no significant influence of caffeine or alcohol intake in the sample.

Mann-Whitney tests revealed no significant differences between groups in age, musical gaming, and sports activity. However, the break-nap group had more participants in the initial years of their study (Median [Mdn]$_{break-nap}$ = 1.0 y; Mdn$_{massed}$ = 3.0 y; Mdn$_{break}$ = 3.0 y; Mdn$_{spaced}$ = 2.0 y). Nonparametric correlations showed no significant relations between years of study and performance on any of the laparoscopic tasks.

One-way analysis of variances showed no significant differences between groups in any of the big 5 personality factors and hours of sleep (the night before training). For quality of sleep (the night before training), the results were mostly comparable, with the exception of a slightly lower sleep quality in the break group compared to the break-nap group (p = 0.03, Mean$_{break}$ = 3.38, standard deviation [SD]$_{break}$ = 0.92; Mean$_{break-nap}$ = 3.92, SD$_{break-nap}$ = 0.76), but without differences for the other 2 groups (Mean$_{massed}$ = 3.78, SD$_{massed}$ = 0.66; Mean$_{spaced}$ = 3.57, SD$_{spaced}$ = 0.83). Further, baseline mental fatigue was significantly higher in the spaced group compared to the break-nap group (p = 0.001, Mean$_{spaced}$ = 22.53, SD$_{spaced}$ = 12.25; Mean$_{break-nap}$ = 12.42, SD$_{break-nap}$ = 8.40), but without differences for the other 2 groups (Mean$_{massed}$ = 18.94, SD$_{massed}$ = 12.10; Mean$_{break}$ = 21.45, SD$_{break}$ = 12.18). Nonparametric correlations showed no significant relations between sleep quality (the night before training) or baseline mental fatigue and performance on any of the laparoscopic tasks.

Results of the activity trackers showed a large difference in the estimated minutes spent asleep between the break and the break-nap group for both of the 2 breaks (break 1: p < 0.001; Mdn$_{break}$ = 0; Mdn$_{break-nap}$ = 10; break 2: p < 0.001; Mdn$_{break}$ = 0; Mdn$_{break-nap}$ = 12). Self-reported minutes asleep in the break-nap group ranged from 0 to 30 (Mean$_{nap1}$ = 5.95; SD$_{nap1}$ = 6.62; Mean$_{nap2}$ = 6.73; SD$_{nap2}$ = 8.22). The activity trackers' measure of minutes spent asleep correlated significantly with the associated self-report measure (r = 0.353, p = 0.032). No significant correlations were found between self-reported sleep the night before training and self-reported (r = −0.178, p = 0.291) and activity trackers' measure (r = −0.148, p = 0.218) of minutes spent asleep during the breaks, although the coefficients were in the direction one would expect.

Main Analysis: Laparoscopic Tasks

For our main analysis, the mean ranks of the Kruskal-Wallis tests are shown in Table 1. If differences between groups were observed, an independent samples Mann-Whitney test was performed to compare individual groups with each other. Figures 1-4 show the results of these tests with asterisks to flag whenever the outcome of these

Mann-Whitney tests proved significant. Figures 1-3 show median scores with 25% to 75% interquartile ratios on each task at each session and serve as an indication of the different learning curves between the training groups. Note that the baseline of the suturing task is lower as only a smaller segment of the task was performed at baseline. This measure was only used for baseline between group analysis, not within-person progress.

At baseline, none of the groups showed significantly different levels of performance on any of laparoscopic tasks. At the end of session 2, we observed a significant difference on the pipe cleaner task between the break and the break-nap group (p = 0.044), whereas the spaced group performed better at the beads task (p = 0.024), the suturing task (p = 0.012), and the total score (p = 0.044) as compared to the massed group. The spaced group also performed significantly better than the break group (p = 0.047) on suturing at the end of the second session. At the end of training, both the spaced and the break-nap group outperformed the massed group on the pipe cleaner (p = 0.044; p = 0.034, respectively) and the total score (p = 0.007; p = 0.009, respectively). The spaced and break-nap group also showed a lower score on the suturing task compared to the massed group, but only borders close to significance for the break-nap group (p = 0.008; p = 0.056, respectively). At retention, both the spaced and break-nap condition outperformed the massed (p = 0.001; 0.001) and the break group (p = 0.005; p = 0.01, respectively) on the pipe cleaner task. On the beads task, the spaced group significantly outperformed both massed (p = 0.017) and break (p = 0.032) groups. On the suturing task, both the spaced and break-nap group outperform the massed group (p = 0.007; p = 0.005, respectively). Both the spaced and break-nap group have a significantly lower total score compared to the massed (p < 0.001; p = 0.001, respectively) and the break group (p = 0.002; p = 0.007, respectively).

One Week Versus One Day Spacing Comparison

In sum, the Mann-Whitney tests revealed substantial differences in completion times on the laparoscopic tasks between the spaced and the massed group. Estimates of effect sizes (r_{rb}) for completion times at the end of the second session, at the end of training, and at retention are displayed in Table 2, along with the effect sizes found in a prior study[3] that used a spacing interval of 1 week. As the retention interval was different from the current study, only end-of-training comparisons have been made.

Sleep and Laparoscopy Measures

To differentiate between the effects of estimated sleep and mere rest (and muted sensory input), we assessed

TABLE 1. Mean Ranks (from Multiple Kruskal-Wallis Tests) for All 4 Training Groups at Baseline, the End of Session II, at the End of Training ($N_{massed} = 40$; $N_{break} = 35$; $N_{break-nap} = 37$; $N_{spaced} = 37$), and at Retention ($N_{massed} = 36$; $N_{break} = 27$; $N_{break-nap} = 35$; $N_{spaced} = 36$)

| | Mean Ranks | | | |
	Massed	Break	Break-nap	Spaced
Measure				
Pipe cleaner$_{baseline}$	76.95	71.37	69.47	81.85
Pipe cleaner$_{end\ of\ session\ II}$	78.95	79.03	62.36	77.66
Pipe cleaner$_{end\ of\ training}$	84.71	81.37	65.58	67.89
Pipe cleaner$_{retention}$	82.99	79.56	55.70	54.44
Beads$_{baseline}$	82.61	70.21	73.78	70.49
Beads$_{end\ of\ session\ II}$	82.85	72.36	75.82	64.07
Beads$_{end\ of\ training}$	84.22	72.89	69.09	71.19
Beads$_{retention}$	74.97	74.63	63.71	56.42
Suturing$_{baseline}$	71.41	76.79	76.42	73.81
Suturing$_{end\ of\ session\ II}$	82.98	79.39	72.03	61.46
Suturing$_{end\ of\ training}$	86.78	72.94	72.15	63.32
Suturing$_{retention}$	81.10	73.00	57.37	59.63
Total z-scores$_{baseline}$	77.63	72.37	73.29	72.30
Totsl z-scores$_{end\ of\ session\ II}$	83.25	76.58	69.00	66.70
Total z-scores$_{end\ of\ training}$	88.51	77.06	64.50	65.05
Total z-scores$_{retention}$	83.42	79.74	54.71	52.64

the relation between self-reported sleep at breaks, self-reported sleep the night before training, the results of the activity trackers at breaks, and performance on the laparoscopic tasks within the subsample of the break-nap group.

Nonparametric correlations showed coefficients ranging from $r = 0.040$ (p = 0.813) to $r = 0.465$ (p = 0.004) for self-reported sleep during the naps and performance on the subsequent laparoscopic tasks. These correlations were most prevalent for the end of training and retention measures, but also present at baseline. However, none of the laparoscopic measures were significantly correlated with the results of the activity trackers' estimate of minutes spent asleep. No significant correlations were found between performance on the laparoscopic tasks and self-reported sleep the night before training.

Rating Scale Mental Effort Analysis

No significant correlations were found between scores on the RSME and estimated time spent asleep at breaks (self-report and activity tracker). At baseline and at the end of session 2, no effects of mental fatigue on laparoscopic performance were found. Fatigue at the end of session 3 did show significant correlations with the pipe cleaner ($r = 0.192$, p = 0.019), suturing ($r = 0.206$, p = 0.012), and total score ($r = 0.219$, p = 0.008) at the end of training and the suturing task at retention ($r = 0.173$, p = 0.046).

The RSME scores are shown in Table 3, along with the results of multiple paired sample t-tests illustrating the progression of the scores for all 4 training groups. A clear reduction in mental fatigue can be observed for the spaced group compared to the 3 other groups.
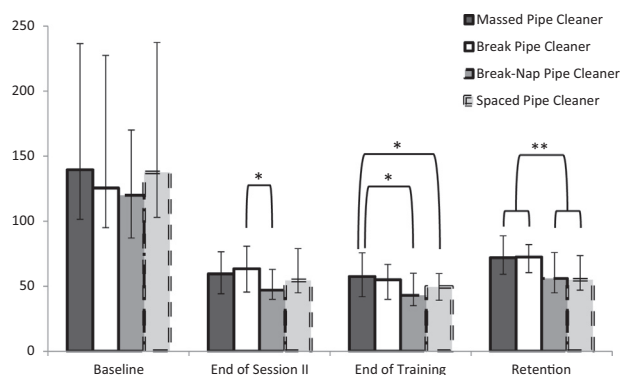

FIGURE 1. Median completion times (in seconds) for the pipe cleaner task at baseline, at the end of the second session, at the end of training, and at retention for all training groups (*p < 0.05; **p < 0.01; ***p < 0.001).
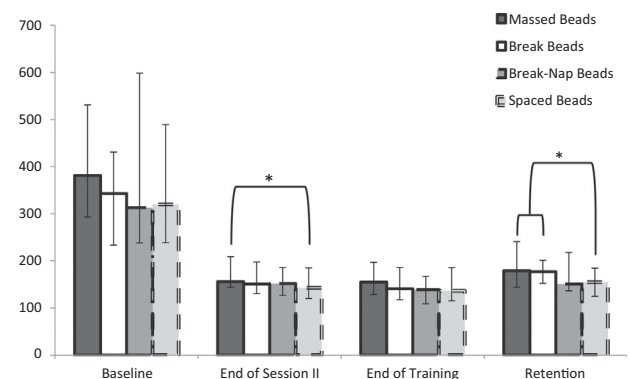

FIGURE 2. Median completion times (in seconds) for the beads task at baseline, at the end of the second session, at the end of training, and at retention for all training groups (*p < 0.05; **p < 0.01; ***p < 0.001).
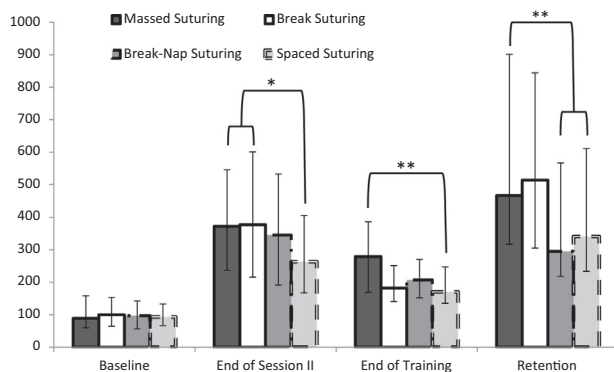
FIGURE 3. Median completion times (in seconds) for the intracorporeal suturing task at baseline, at the end of the second session, at the end of training, and at retention for all training groups (*p < 0.05; **p < 0.01; ***p < 0.001). Please note that only a segment of the suturing task was performed at baseline (discussed in Method section), hence the lower completion times.

DISCUSSION

The findings in this study further nuanced the current theory regarding the spacing effect as it applies to laparoscopic surgical training. We found that spacing 3 training sessions across 3 consecutive days is advantageous compared to a massed 1-day schedule, even if that schedule accounts for substantial amounts of breaks in training sessions.

Alternatively, we found that including powernaps in between the training sessions on a 1-day schedule, can enhance long-term retention, an option which may be more beneficial in terms of logistics when organizing training events, when time constraints are a bigger motive.

THE EFFECTS OF NAPS

One of the more striking findings was that within the break-nap group, participants who reported spending more time asleep during the naps had worse performance on the laparoscopic tasks at the end of training and retention. This
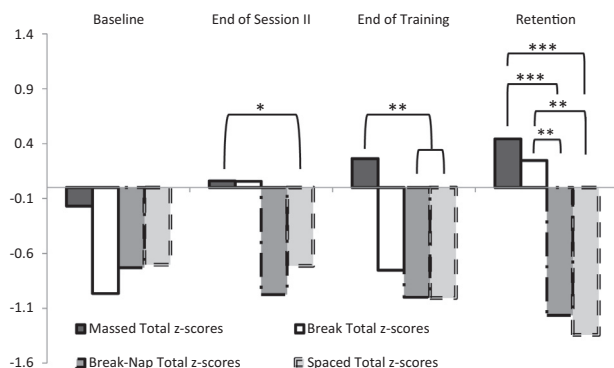

FIGURE 4. Median total z-scores at baseline, at the end of the second session, at the end of training, and at retention for all training groups (NS, nonsignificant; *p < 0.05; **p < 0.01; ***p < 0.001).

seems counterintuitive as the break-nap group as a whole shows a pattern of better performance compared to the break group, the most similar condition that does not allow sleep during the break.

A possible explanation is that the lower performance was caused by sleep inertia, a time period of impaired performance, and disorientation as a person transitions from sleep to wakefulness.[15] We accounted for this possibility in the schedule of our sessions, as all performance measurements were done directly before the naps, with the naps being followed by a brief period to fill out a questionnaire and resuming training on a basic task in the next session. Still, sleep inertia may have effected trainees while practicing during the following session, or the inertia effects may not have fully wore off by the time of the next measurement.[16] The occurrence of sleep inertia can differ between contexts. The recommendation on the duration for daytime naps is to keep them less than 30 minutes, as longer naps are associated with lower productivity and increased sleep inertia.[17] However, when employees work in night shifts or have prolonged working hours, longer naps (40-60 min) show more performance benefits.[18] Additionally, effects on motor performance can be influenced by whether a trainee habitually takes naps or not.[19]

Another explanation is that participants who slept (longer) within the break-nap group did so because they were more sleep deprived and sleepy to begin with, and this led to a suboptimal state for learning throughout the day of training,[15] while the naps were not sufficient to decrease sleep deprivation and sleepiness. Contrary to this explanation, there was no association between sleep quantity and sleep quality on the night before the training and performance on the laparoscopic tasks (both within the break-nap group and the overall sample).

An alternative explanation is that within the break-nap group, trainees with a lower aptitude for learning laparoscopic skills experienced more arousal due to their difficulties in learning the tasks and reported being more fatigued and having spent more time asleep during the breaks as a misattribution of their arousal levels.[20] In any case, the findings should be interpreted cautiously as no correlations were found between the performance on the laparoscopic tasks and estimated sleep derived from the activity monitoring devices. Regardless of whether sleep during naps deteriorates performance, the results do indicate that trainees benefit from a period of rest while they are inactive, lying down, and while their sensory input is muted as compared to a traditional break.

THE EFFECTS OF FATIGUE AND CONSOLIDATION

Even though participants in the massed group reported experiencing significantly more fatigue compared to the spaced

TABLE 2. Effect Sizes (r_{rb}) of the Mann-Whitney Tests for the 3 Training Tasks

	Break-nap|break	Break-nap|massed	Spacing|break	Spacing|massed	Spacing|massed[3]
Pipe cleaner task					
End of session II	0.20				
End of training		0.21		0.19	0.73
Retention	0.29	0.35	0.32	0.37	
Beads task					
End of session II				0.23	
End of training					0.65
Retention			0.23	0.25	
Suturing task					
End of session II			0.20	0.26	
End of training				0.27	0.58
Retention		0.31		0.29	

Effect sizes are only mentioned when tests were significant and when a comparison was viable (for the previous spacing study).

group, we observed weak correlations between fatigue and performance on the laparoscopic tasks in the sample, which suggests a small influence of fatigue contributing to the spacing effect. It is also worth considering that fatigue may have a higher impact on learning efficiency when a higher level of fatigue is reached, but that this threshold was simply never reached using the current design of training.

Learning of motor skills largely occurs in the timeframe following training,[9] when there is opportunity for consolidation of the training stimuli. Consolidation occurs during "off-line" periods simply as time passes, but is enhanced with a short nap or during overnight sleep,[8,21,22] although this may depend on the nature of the task.[21] Typically, effects of overnight sleep consolidation are enhanced for more difficult tasks[23] and when a break is planned early in the training schedule rather than later.[24] Other authors have suggested that sleep effects can be attributed to the type of design used (sleep deprivation control groups that tend to impair performance), and that mere periods of rest show similar improvements as sleep.[25] Regardless of whether sleep enhances consolidation or whether sleep deprivation impairs performance (or both), one can conclude that sleeping sufficiently is beneficial after training. The results of the current study also suggest that night sleep is more beneficial in between training sessions than just after the entire training, as the spaced group (who had 2 nights of sleep in between training sessions) had the best performance at the end of training and at retention, followed by break-nap group, the break group, and the massed group, respectively.

When comparing the 2 spaced groups (the one in the current study and the one from the previous study[3]), we observed higher effect sizes for the intervention with the bigger spacing interval (1 wk). Future research should investigate what the optimal spacing interval is, although this is likely influenced by the desired retention interval.[26] Finally, trainers ought to be cautious in the design of their training, as consolidation can be disrupted during the acquisition (active training) phase when multiple conflicting training conditions are interleaved during training,[27,28] for instance when multiple tasks are trained that are too dissimilar or when they have opposing motor patterns.

LIMITATIONS AND CONCLUSION

One of the limitations of the study is that we did not include any measures of self-efficacy in our design. In the introduction, we noted that spacing can enhance more accurate appraisal of the skill level of a trainee (by both trainee and instructors), and with the current design we were unable to estimate the extent to which this factor influenced the effectiveness of spacing. Furthermore, we used a different retention interval from the previous study,[3] so we only have comparisons of effect sizes for the end of training and not retention. The current study only used completion times, so no generalizations to other performance outcome measures (accuracy, instrument path length, force, etc.) can be made from these findings.

TABLE 3. Means Rating Scale Mental Effort Scores With Paired Samples t-Tests Assessing the Progression in Mental Fatigue for All 4 Training Groups at Baseline, the End of Session II, and at the End of Training

	Mean$_{massed}$	Mean$_{break}$	Mean$_{break-nap}$	Mean$_{spaced}$
Measure				
RSME$_{baseline}$	18.94	21.45	12.42	22.53
RSME$_{end\ of\ session\ II}$	33.81↑***	40.67↑***	35.43↑***	25.15≈
RSME$_{end\ of\ training}$	50.61↑***	48.50≈	48.30↑***	27.55≈

≈ For no significant difference; ↓/↑ for significant increase/decrease, *p < 0.05; **p < 0.01; ***p < 0.001

Besides, we used self-reports and activity monitoring devices to measure sleep, but these only serve as an approximation of time spent asleep. Future studies could make use of polysomnography to provide a more valid estimate on the influence of sleep consolidation on skill acquisition and retention of laparoscopic motor skills.

To answer the initial research question, we conclude that consolidation has the biggest influence in the effectiveness of spacing. Mitigating mental fatigue by inclusion of breaks and providing small periods of sensory muting are beneficial, but play a smaller role. When applying powernaps in training, one ought to be cautious for sleep inertia, and it is recommended to schedule short time frame (less than 30 min)[17] with ample rest time after waking to allow for waning of potential sleep inertia. Another way of preventing a poor learning state and poor consolidation posttraining is to instruct trainees to sleep sufficiently before and after each training session. Finally, when comparing the effect sizes of this study to the ones found previously,[3] we found that spacing over 3 weeks leads to better learning efficiency and retention than spacing over 3 days. We urge training designers to take into consideration factors such as consolidation, fatigue, sleep, boredom, and self-efficacy when developing a surgical curriculum.

ACKNOWLEDGMENTS

REFERENCES

1. Moulton Carol-Anne, Dubrowski Adam, MacRae Helen, Graham Brent, Grober Ethan, Reznick Richard. Teaching surgical skills: what kind of practice makes perfect?: A randomized, controlled trial *Ann Surg*. 2006;244(3):400.

2. Gallagher Anthony G, Jordan-Black Julie Anne, O'Sullivan Gerald C. Prospective, randomized assessment of the acquisition, maintenance, and loss of laparoscopic skills. *Ann Surg*. 2012;256(2):387-393.

3. Spruit Edward N, Band Guido PH, Hamming Jaap F. Increasing efficiency of surgical training: effects of spacing practice on skill acquisition and retention in laparoscopy training. *Surg Endosc*. 2015;29(8): 2235-2243.

4. Kahol Kanav, Leyba Mario J, Deka Mary, et al. Effect of fatigue on psychomotor and cognitive skills. *Am J Surg*. 2008;195(2):195-204.

5. Bjork, Robert A. Assessing our own competence: heuristics and illusions. In: Gopher Daniel, Koriat Asher, eds. Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application; 1999.

6. Wang Zuowei, Zhou Renlai, Shah Priti. Spaced cognitive training promotes training transfer. *Front Hum Neurosci*. 2014;8:1-8, article 217.

7. Stickgold Robert. Sleep-dependent memory consolidation. *Nature*. 2005;437(7063):1272-1278.

8. Brashers-Krug Thomas, Shadmehr Reza, Bizzi Emilio. Consolidation in human motor memory. *Nature*. 1996;382(6588):252-255.

9. Walker Matthew P, Brakefield Tiffany, Morgan Alexandra, Hobson J Allan, Stickgold Robert. Practice with sleep makes perfect: sleep-dependent motor skill learning. *Neuron*. 2002;35(1):205-211.

10. Korman Maria, Doyon Julien, Doljansky Julia, Carrier Julie, Dagan Yaron, Karni Avi. Daytime sleep condenses the time course of motor memory consolidation. *Nat Neurosci*. 2007;10(9):1206-1213.

11. Kolkman Wendela, van de Put Mathijs AJ, Wolterbeek Ron, Trimbos J Baptist MZ, Jansen Frank W. Laparoscopic skills simulator: construct validity and establishment of performance standards for residency training. *Gynecol Surg*. 2008;5(2):109-114.

12. Gosling Samuel D, Rentfrow Peter J, Swann William B. A very brief measure of the Big-Five personality domains. *J Res Pers*. 2003;37(6):504-528.

13. Zijlstra Ferdinand RH. Efficiency in Work Behaviour: A Design Approach For Modern Tools. Delft, The Netherlands: TU Delft, Delft University of Technology; 1993.

14. Montgomery-Downs Hawley E, Insana Salvatore P, Bond Jonathan A. Movement toward a novel activity monitoring device. *Sleep Breath*. 2012;16(3):913-917.

15. Miccoli Laura, Versace Francesco, Koterle Sara, Cavallero Corrado. Comparing sleep-loss sleepiness and sleep inertia: lapses make the difference. *Chronobiol Int*. 2008;25(5):725-744.

16. Burke Tina M, Scheer Frank AJL, Ronda Joseph M, Czeisler Charles A, Wright JR, Kenneth P. Sleep inertia, sleep homeostatic and circadian influences on higher-order cognitive functions. *J Sleep Res*. 2015.

17. Dhand Rajiv, Sohal Harjyot. Good sleep, bad sleep! The role of daytime naps in healthy adults. *Curr Opin Pulm Med*. 2006;12(6):379-382.

18. Mulrine Hannah M, Signal T Leigh, van den Berg Margo J, Gander Philippa J. Post-sleep inertia performance benefits of longer naps in simulated night-work and extended operations. *Chronobiol Int.* 2012;29(9):1249-1257.

19. Milner Catherine E, Fogel Stuart M, Cote Kimberly A. Habitual napping moderates motor performance improvements following a short daytime nap. *Biol Psychol.* 2006;73(2):141-156.

20. Cotton John L. A review of research on Schachter's theory of emotion and the misattribution of arousal. *Eur J Soc Psychol.* 1981;11(4):365-397.

21. Doyon Julien, Korman Maria, Morin Amélie, et al. Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res.* 2009;195(1):15-26.

22. Korman Maria, Raz Naftali, Flash Tamar, Karni Avi. Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proc Natl Acad Sci.* 2003;100(21):12492-12497.

23. Kuriyama Kenichi, Stickgold Robert, Walker Matthew P. Sleep-dependent learning and motor-skill complexity. *Learn Mem.* 2004;11(6):705-713.

24. Duke Robert A, Allen Sarah E, Cash Carla D, Simmons Amy L. Effects of early and late rest breaks during training on overnight memory consolidation of a keyboard melody. *Ann N Y Acad Sci.* 2009;1169 (1):169-172.

25. Rieth Cory A, Cai Denise J, McDevitt Elizabeth A, Mednick Sara C. The role of sleep and practice in implicit and explicit motor learning. *Behav Brain Res.* 2010;214(2):470-474.

26. McDaniel Mark. Put the SPRINT in Knowledge Training: Training with SPacing, Retrieval, and INTerleaving. Healy Alice F, Bourne Lyle Eugene, editors. Training Cognition. New York: Psychology Press, 2012.

27. Banai Karen, Ortiz Jeanette A, Oppenheimer Jason D, Wright Beverly A. Learning two things at once: differential constraints on the acquisition and consolidation of perceptual learning. *Neuroscience.* 2010;165 (2):436-444.

28. Spruit Edward N, Kleijweg Luca, Band Guido PH, Hamming Jaap F. Varied practice in laparoscopy training: beneficial learning stimulation or cognitive overload? *Front Psychol.* 2016;7:685.

SUPPLEMENTARY MATERIAL

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.jsurg.2016.11.003.