Comparing Technical Dexterity of Sleep-Deprived Versus Intoxicated Surgeons

Fariba Mohtashami, MD, Allison Thiele, MD, PhD, Erwin Karreman, PhD, John Thiel MD

ABSTRACT

Background: The evidence on the effect of sleep deprivation on the cognitive and motor skills of physicians in training is sparse and conflicting, and the evidence is nonexistent on surgeons in practice. Work-hour limitations based on these data have contributed to challenges in the quality of surgical education under the apprentice model, and as a result there is an increasing focus on competency-based education. Whereas the effects of alcohol intoxication on psychometric performance are well studied in many professions, the effects on performance in surgery are not well documented. To study the effects of sleep deprivation on the surgical performance of surgeons, we compared simulated the laparoscopic skills of staff gynecologists “under 2 conditions”: sleep deprivation and ethanol intoxication. We hypothesized that the performance of unconsciously competent surgeons does not deteriorate postcall as it does under the influence of alcohol.

Methods: Nine experienced staff gynecologists performed 3 laparoscopic tasks in increasing order of difficulty (cup drop, rope passing, pegboard exchange) on a box trainer while sleep deprived (<3 hours in 24 hours) and subsequently when legally intoxicated (>0.08 mg/mL blood alcohol concentration). Three expert laparoscopic surgeons scored the anonymous clips online using Global Objective Assessment of Laparoscopic Skills criteria: depth perception, bimanual dexterity, and efficiency. Data were analyzed by a mixed-design analysis of variance.

Results: There were large differences in mean performance between the tasks. With increasing task difficulty, mean scores became significantly (P < .05) poorer. For the easy tasks, the scores for sleep-deprived and intoxicated participants were similar for all variables except time. Surprisingly, participants took less time to complete the easy tasks when intoxicated. However, the most difficult task took less time but was performed significantly worse compared with being sleep deprived. Notably, the evaluators did not recognize a lack of competence for the easier tasks when intoxicated; incompetence surfaced only in the most difficult task.

Conclusions: Being intoxicated hinders the performance of more difficult simulated laparoscopic tasks than being sleep deprived, yet surgeons were faster and performed better on simple tasks when intoxicated.

Key Words: Surgical education, Sleep deprivation, Ethanol intoxication, Laparoscopic performance, Simulation.

INTRODUCTION

The question of work hours and competence has implications for practicing surgeons. The current evidence on the effect of sleep deprivation comes from studies performed on physicians in training rather than surgeons in practice. However, the 2 groups are not equivalent, and results may not translate. According to Fitts and Posner's theory, manual skills are acquired in 3 stages: cognitive, associative, and autonomous. An individual theoretically understands a task in the cognitive stage and then transfers this knowledge to task performance by associating cognitive elements with musculoskeletal maneuvers in the associative stage. The autonomous stage is reached after repetition, and when the task completion requires minimal demands on attention, human cognitive capacity can be used for additional activities. Residents are most likely in the associative stage of motor learning for technical skills, and the results of such studies may not be reliably extrapolated to physicians in practice, who are by definition in the autonomous stage.

The evidence on the effect of work hours and fatigue on the acquisition and performance of technical surgical skills is conflicting. Some studies have shown a negative impact of sleep deprivation on residents' performance, cognition, and mood, whereas others have found no ef-
fect of fatigue on psychomotor skills: Eastridge et al. and Taffinder et al. showed deteriorated simulated laparoscopic performance in surgical residents after a night on call. The residents made more errors and took more time to complete the tasks postcall. Other reports found impaired language and math skills, deteriorated electrocardiographic interpretation, and decreased mood, concentration, and cognitive performance in postcall residents, but Clancy et al. did not find any significant decrease in motor skills of residents postcall. A prospective study looked at surgical outcomes in 4,047 cardiac surgical procedures when resident surgeons had 0 to 3, 3 to 6, and >6 hours of sleep the night before the procedures and found no difference in mortality rates and major complications. A meta-analysis of 19 studies concluded that sleep deprivation affects mood the most and motor skills the least. Finally, investigators have even found improved performance in simulated laparoscopic tasks and concluded that night call does not impair the learning of laparoscopic skills.

To describe the effects of fatigue on performance, some have compared sleep deprivation to alcohol g/dL intoxication. Alcohol is a depressant that reduces cognitive and motor activity. The legal limit of blood alcohol concentration (BAC) for driving in most countries is 0.08, as judgment, self-control, caution, reasoning, memory, and motor coordination decrease between 0.08 and 0.12 BAC. Loss of consciousness can occur at 0.30 BAC, and at 0.40 BAC, respiratory arrest can lead to death. Alcohol’s effects have been tested against standardized psychometric tests such as the Trail Making Test and Occupational Safety Performance Assessment Technology, which are frequently used in clinical neuropsychology to measure motor performance. On the basis of laboratory, simulator, and on-road measures of speed and accuracy, the effects of alcohol provide a benchmark against which motor and cognitive performance can be measured. Dawson and Reid compared the effects of sustained wakefulness with alcohol intoxication on cognitive and psychomotor performance. They showed that the performance decrement for each hour of wakefulness after 10 to 26 hours was equivalent to the performance decrement observed with a 0.04% rise in BAC. However, no study has looked at the effects of alcohol on the motor skills required to be a competent surgeon.

Our study was designed to test surgical competence in sleep-deprived and alcohol-intoxicated surgeons. We hypothesized that sleep deprivation would not affect the technical dexterity of autonomous surgeons but that performance would deteriorate at 0.08 BAC. We further hypothesized that blinded evaluators would not distinguish the difference between rested and postcall performance, but at ≥0.08 BAC, the deterioration in performance would become evident to the evaluators.

MATERIALS AND METHODS

The study was performed in a tertiary care hospital in Regina Qu'Appelle Health Region at the University of Saskatchewan. Approval for the study was obtained from the Regina Qu'Appelle Health Region Research Ethics Board. Invitation letters were sent to each gynecologist in Regina who does laparoscopic surgery. The study consisted of performing 3 laparoscopic tasks in increasing order of difficulty on a conventional box trainer, and the performance was recorded.

Participants had the chance for individual instruction in the 3 tasks in the technical skills laboratory. The scoring system and definition of errors were explained. Each surgeon repeated the tasks in the same order after a night on call with <3 hours of sleep in 24 h and subsequently under the influence of alcohol. To minimize the effect of circadian variation in cognitive psychomotor performance of the subjects, the postcall tests were completed between 8 and 11 AM.

In the final assessment, the surgeons were offered a selection of alcoholic beverages, and their BAC was measured periodically by a standard calibrated breath analysis (S80 Pro Breathalyzer, BACtrack, San Francisco, California) that was accurate to 0.005% BAC. All participants were regular social drinkers with an approximate 4 to 8 drinks per week and were familiar with the effects of alcohol. For this reason, a placebo condition was not included in the study design.

The subjects stopped drinking 15 minutes before BAC was measured. Once BAC levels ≥ 0.08 were reached, each subject performed the tasks in the same increasing order of difficulty, and the performance was recorded. We selected 0.08 as the minimum BAC as proof of impairment. Because the subjects were familiar with the tasks, the total time between the first at home third task was not more than 10 minutes. Safe transportation was organized for the participants after the completion of the study.

The video clips were posted on a secure Web site and scored by 3 independent reviewers. The evaluators were blinded as to the surgeon and whether the condition was rested, postcall, or intoxicated. The scoring was based on number of errors, time to complete the task, efficiency, depth perception, and bimanual dexterity. Each clip was
also scored for the overall performance along a 5-point scale and rated either competent or incompetent.

**Simulated Laparoscopy Tasks**

**Task 1: Cup Drop.** The participants used graspers to transfer 10 beads and drop them one by one into a cylindrical container with each dominant and nondominant hand. Time was recorded when both graspers appeared on the screen and stopped when the last bead was dropped in the container and the graspers were withdrawn. Errors were calculated as the number of beads dropped. The cup-drop task can be viewed on YouTube at https://www.youtube.com/user/faribamohtashami.

**Task 2: Rope Passing.** Participants used 2 graspers to manipulate a 40-cm rope by passing it from one hand to the other in 1 direction. They grasped the rope on the black-mark targets that were placed at 2.5-cm intervals. The task was then repeated in the opposite direction. Time was recorded from the time 2 graspers appeared on the screen until the rope was dropped after the end of the rope was reached and the graspers were withdrawn. Errors were defined as grasping the rope outside black marks and dropping rope. The rope-passing task can be viewed on YouTube at https://www.youtube.com/user/faribamohtashami.

**Task 3: Pegboard Exchange.** Participants used 2 graspers to pick up 10 triangular plastic pegs one by one with one hand, transfer it to the other hand in the air, and place it on the board. The subjects had the option of using either the dominant or nondominant hand to pick up the pegs, but the same order was followed through the test. The time to complete the task was measured from the time both graspers appeared on the screen until the last peg was placed on the board and the graspers were withdrawn. Errors were counted for dropping the pegs or failing to transfer the peg in the air by using the board to facilitate the task. The pegboard-exchange task can be viewed on YouTube at https://www.youtube.com/user/faribamohtashami.

**Statistical Analysis**

Each subject was compared to himself or herself and among the group. A $2 \times 3$ repeated-measures factorial analysis-of-variance design was conducted to analyze the data. One factor in the analysis consisted of the 3 tasks, while the other factor was the state of the participant (sleep deprived or legally intoxicated). Main effects and an interaction effect were analyzed. For this analysis, the rather small sample size was compensated for by the fact that each participant’s data were represented in each cell of the analysis-of-variance design.

**RESULTS**

There were large differences in the overall performance between tasks (Figure 1). With increasing task difficulty, mean scores became significantly poorer ($P < .001$). For the easier tasks of cup drop (task 1) and rope passing (task 2), the scores for sleep deprivation and intoxication differed little for all variables except time (Figure 2). Surprisingly, participants took less time to complete the easy tasks (tasks 1 and 2) when intoxicated. However, for pegboard exchange, the most difficult task, the participants took less time but performed significantly worse when intoxicated compared with sleep deprived, as determined by overall performance.

Other individual variables measured were errors (Figure 3), efficiency (Figure 4), depth perception (Figure 5), and bimanual dexterity (Figure 6). For every variable measured, the difference between tasks was significantly different (Table 1). Participants did significantly worse the harder the task they had to complete. However, for none of the outcome variables did the state of the participants differ significantly. The effect size in this case was $r = 0.62$, which is considered a large effect. In other words, participants were faster in the intoxicated state (Table 1).

![Figure 1](https://www.youtube.com/user/faribamohtashami)
Most interesting are the results that describe the interaction effect between the different levels of task and state. As can be seen from all variables except time (Figure 2), scores for sleep deprivation and intoxication were similar for tasks 1 and 2 (Figures 4–6), with scores for intoxication being slightly better. Then for task 3, scores for intoxication declined more than for sleep deprivation and were worse in all these cases (Figures 3–6). Although none of the overall interaction effects was significant, contrasts between the individual levels of task and state did show some significant effects. Furthermore, when tasks 2 and 3 were compared, effect sizes for all variables except time indicated that large effects were present (ranging from 0.51 to 0.69, while the effect size for time is considered small at 0.15). A similar trend was found when tasks 1 and 3 were compared (Table 1). Overall, these results suggest that although the differences between intoxication and sleep deprivation were minimal for tasks 1
and 2, the decline in task 3 scores for intoxication was larger than for sleep deprivation. Hence, being intoxicated is more of a disadvantage than being sleep deprived.

As shown in Table 2, there was no difference between the number of times the performance was rated “competent” for rested compared with postcall for all 3 tasks. This suggests that participants performed equally well in both states. When comparing rested and postcall (competent) with intoxicated (incompetent), evaluators rated intoxicated performance “competent” for tasks 1 and 2 more often than rested and postcall combined. Thus, evaluators were unable to recognize the lack of competence for the simple tasks, and the correct recognition of incompetence was significant only for the difficult task (Table 3).

DISCUSSION

Few would argue that surgical competence requires a combination of technical and nontechnical skills, yet how these are acquired and measured remains unclear. Competence requires not only the ability to carry out a task but also the behavior, personal characteristics, and attributes to know when is the best time to perform that task. Technical performance is by far the hallmark of surgeons, and our study focused on technical dexterity. It has been speculated that surgeons who are sleep deprived are not technically competent. As it would be difficult to identify and then study the performance of incompetent surgeons, we deconstructed technical competence using the standard pharmacologic effect of alcohol to study its effects and compare them with those of sleep deprivation.

Our study demonstrated that sleep deprivation does not impair the technical dexterity of gynecologic surgeons who are “unconsciously competent” in performing basic laparoscopic tasks. The evaluators did not record any significant difference between rested and postcall performance for easy and difficult tasks and rated both as competent. Surprisingly, however, the participants performed the easy tasks even better when intoxicated: they needed less time; had better depth perception, bimanual dexterity, and efficiency; and scored higher on overall performance. The evaluators were not able to recognize the intoxicated state when simple tasks were performed. Performance did deteriorate when the surgeons were intoxicated only for the difficult task, with the participants making more errors and having worse depth perception, bimanual dexterity, and efficiency and poorer overall performance. It is likely that intoxicated surgeons lack recognition of their own abilities, lose caution, and take miscalculated risks. Exercising poorer judgment leads to poorer performance. Despite the frequent comparison of sleep deprivation and alcohol intoxication, the evaluators recognized the lack of competence in the intoxicated surgeons only during the difficult task.

Our study shows that competent surgeons retain their technical dexterity for simple tasks when sleep deprived or when intoxicated. More important, they are able to retain their competence for complex tasks when sleep deprived, which is highly relevant in scheduling type and load of work postcall. However, by no means do we imply that surgeons can attend their easy days intoxicated. Our data suggest that decisions regarding surgeons’ work hours should be based on studies involving competent surgeons rather than on surgical learners or other professions. Limitations on hours of work based on such comparisons could result in erroneous conclusions, reduced surgical output, and longer surgical waits for patients.

The contributing factors to reduced surgical experience for graduates of obstetrics and gynecology residency programs have been extensively published over the past 20 years. Critical factors required to produce competent surgeons under apprenticeship are a high volume of cases and long work hours leading to repetitive exposure with the same mentor. The inevitable changes to the practice of gynecology and limits on resident work hours have significantly compromised this teaching model. In July 2003 the Accreditation Council for Graduate Medical Education imposed regulations that limited residents’ duty hours to

![Figure 6. The effects of sleep deprivation and alcohol intoxication on bimanual dexterity in performing simulated laparoscopic tasks. Task 1, cup drop; task 2, rope passing; task 3, pegboard exchange. The bimanual dexterity was scored along a 5-point scale for 9 subjects and the mean plotted.](image)
80 hours per week. The work hours for European medical trainees are limited to 48 hours per week.19 There are similar legislation in other safety-sensitive industries such as aviation and commercial truck driving. The Federal Aviation Administration limits the maximal allowable flight time for a single pilot voyage to 8 hours,20 and the Federal Motor Carrier Safety Administration prohibits driving more than 11 hours per day and 60 hours per week.21 Airline pilots and surgeons are similar in that both their professions require the acquisition and mastery of technical skills and knowledge of how to deal with potentially life-threatening conditions. The safety of the clients directly depends on the performance and level of proficiency of the operator.22 Although there have been changes made to the allowable work hours for residents, there is no legislation to limit the work hours for surgeons and physicians in practice. A comprehensive review of the literature related to restricted duty hours in surgical residents showed negative impacts on patient outcomes and performance on certification examinations. Restricted duty hours were not consistently associated with improvements in residents’ well-being.23 Inadequate surgical proficiency contributes to poor surgical outcomes,24 and resulting complications may be compounded by the relative inexperience of the surgeon.25,26 However, measurement of competence and how to apply this to individual surgeons is a difficult yet critical task and requires the ability to accurately distinguish among different levels of performance. This is problematic, as an inability to measure means an inability to assess. Our study showed that correct assessment could be reliably made only when assessing intoxicated-incompetent surgeons performing difficult tasks. There is a significant overlap between competent and incompetent performance on simple tasks, as “everyone can sail in a calm sea.” Thus, for an evaluation of competence when challenged, simple tasks must be excluded from assessments.

Any curriculum developed in this ambiguous environment is potentially at risk for being added on to an already

### Table 1.

Significance Testing Between Categories of States (Sleep Deprivation and Alcohol Intoxication) and Task

| Variable          | Time | Errors | Efficiency | Depth Perception | Bimanual Dexterity | Overall Performance |
|-------------------|------|--------|------------|------------------|--------------------|---------------------|
| State (P)         | .053 | .202   | .494       | .511             | .598               | .563                |
| Task (P)          | <.001| <.001  | <.001      | <.001            | <.001              | <.001              |
| Interaction (P)   | .741 | .175   | .075       | .098             | .057               | .074                |
| Task 1 vs task 3 | .771 | .374   | .091       | .273             | .010               | .043                |
| Task 2 vs task 3 | .687 | .136   | .029       | .027             | .058               | .056                |
| Interaction (r)   |      |        |            |                  |                    |                     |
| Task 1 vs task 3 | 0.11 | 0.32   | 0.56       | 0.38             | 0.76               | 0.65                |
| Task 2 vs task 3 | 0.15 | 0.51   | 0.69       | 0.69             | 0.62               | 0.62                |

Differences between categories of states (sleep deprivation and alcohol intoxication) and tasks (P values are reported). The variables in the first column refer to the main effect for the state of the participant and the task he or she was performing. The interaction between these two variables is listed as indicated. P represents significance level, while r represents effect size. Task 1, cup drop; task 2, rope passing; task 3, pegboard exchange.
overloaded curriculum of obstetrics and gynecology training programs and aggravating the existing dysfunction in the surgical education of residents. The need to understand the components of surgical competence and to develop a valid assessment tool is paramount for successful implementation of competency-based education. The profession also needs to reach consensus on the definition of competence, “acceptable” level of competence, and proficiency for gynecologic surgeons before being able to guide and design competency-based curricula for novices.

We gratefully acknowledge the help in scoring the performance video clips online by Drs H.W.R. Schreuder (University Medical Centre Utrecht) and M. Alkayki (University of Saskatchewan).

References:
1. Fitts PM, Posner MI. Human performance. Basic concepts in psychology series. Belmont, CA: Brooks/Cole; 1967.
2. Logan GD. Automaticity, resources, and memory: theoretical controversies and practical implications. Hum Factors. 1988;30:583–598.
3. Eastridge BJ, Hamilton EC, O’Keefe GE, et al. Effect of sleep deprivation on the performance of simulated laparoscopic surgical skill. Am J Surg. 2005;186:169–174.
4. Taffinder NJ, McManus IC, Gul Y, et al. Effect of sleep deprivation on surgeons’ dexterity on laparoscopy simulator. Lancet. 1998;352:1191.
5. Hawkins MR, Vichick DA, Silshy HD, et al. Sleep and nutritional deprivation and performance of house officers. J Med Educ. 1985;60:530–535.
6. Friedman RC, Bigger JT, Kornfeld DS. The intern and sleep loss. N Engl J Med. 1971;285:201–203.
7. Dru M, Bruge P, Benoit O, et al. Overnight duty impairs behaviour, awake activity and sleep in medical doctors. Eur J Emerg Med. 2007;14(4):199–203.
8. Pilcher JJ, Huffcutt AI. Effects of sleep deprivation on performance: a meta-analysis. Sleep. 1996;19(4):318–326.
9. Koslowsky M, Babkoff H. Meta-analysis of the relationship between total sleep deprivation and performance. Chronobiol Int. 1992;9:132–136.
10. Clancy K, Brady P, McHugh S, et al. Assessing the functional performance of post-call hospital doctors using a Nintendo Wii. Ir Med J. 2011;104(6):171–173.
11. Chu MWA, Stitt LW, Fox SA, et al. Prospective evaluation on consultant surgeon sleep deprivation and outcomes in more than 4000 consecutive cardiac surgical procedures. Arch Surg. 2011;146(9):1080–1085.
12. DeMaria EJ, McBride CL, Broderick TJ, Kaplan BJ. Night call does not impair learning of laparoscopic skills. Surg Innov. 2005;12(2):145–149.
13. Vonghia L, Leggio L, Ferrulli A, Bertini M, Gasbarrini G, Addolorato G; Alcoholism Treatment Study Group. Acute alcohol intoxication. Eur J Intern Med. 2008 Dec;19(8):561–7. doi: 10.1016/j.ejim.2007.06.033. Epub 2008 Apr 2.
14. Lehmann S, Martus P, Little Elk S, et al. Impact of sleep deprivation on medium-term psychomotor and cognitive performance of surgeons: prospective cross-over study with a virtual surgery simulator and psychometric tests. Surgery. 2010;147(2):246–254.
15. Dawson D, Reid K. Fatigue, alcohol and performance impairment. Nature. 1997;388:255.
16. Holvat P, Sleet D, Smith I. Alcohol and driving: is the 0.05% blood alcohol concentration limit justified? Drug Alcohol Rev. 1991;10:151–166.
17. Sorosky J, Anderson B. Surgical experiences and training of residents: perspective of experienced gynecologic oncologists. Gynecol Oncol. 1999;75:222–223.
18. Rathat G, Hoa D, Gagnayre R, Hoffet M, Mares P. Surgical training of trainees, specialists in obstetrics and gynecology: results from a national electronic survey. J Gynecol Obstet Biol Reprod (Paris). 2008;37:672–684.
19. Olson E, Drage LA, Auger RR. Sleep deprivation, physician performance, and patient safety. Chest. 2009;136:1389–1396.
20. Federal Aviation Administration. 14 CFR Part 91. Available at: http://ecfr.gpoaccess.gov/cgi. Accessed February 11, 2009.
21. Department of Transportation, Federal Motor Carrier Safety Administration. 49 CFR 385 and 395. Available at: http://www.fmcsa.dot.gov/rulesregulations/administration/rulemakings. Accessed February 11, 2009.
22. Moen MD, Moen RO, Moen OL. Teaching and evaluating surgical skills: applying the aviation model to gynecologic surgery. J Pelvic Med Surg. 2007;13:107–12.
23. Ahmed N, Devitt KS, Keshet I, et al. A systematic review of the effects of resident duty hour restrictions in surgery. Ann Surg. 2014;259:1041–1053.
24. The Bristol Royal Infirmary Inquiry. The inquiry into the management of care of children receiving complex heart surgery at the Bristol Royal Infirmary. Available at: http://www.bristol-inquiry.org.uk. Accessed July 9, 2007.
25. Meranay AM, Samee AA, Gill IS. Vascular and bowel complications during retroperitoneal laparoscopic surgery. J Urol. 2004;168:1941–1944.
26. Simon SD, Castle EP, Ferrigni RG, et al. Complications of laparoscopic nephrectomy: the Mayo Clinic experience. J Urol. 2004;171:1447–1450.