

Surgical Education

Examination stress leads to improvements on fundamental technical skills for surgery

Vicki LeBlanc, PhD^{a,b,*}, Sarah I. Woodrow, MD, MEd^c, Ravi Sidhu, MD, MEd, FRCSC^d, Adam Dubrowski, PhD^{a,c}

^aWilson Centre, Faculty of Medicine, University of Toronto, Ontario, Canada; ^bORNGE Transport Medicine, Toronto, Ontario, Canada; ^cDepartment of Surgery, University of Toronto, Ontario, Canada; ^dDepartment of Surgery, University of British Columbia, British Columbia, Canada

KEYWORDS:

Stress;
Examinations;
Performance;
Technical skills

Abstract

BACKGROUND: This study was conducted to assess the effects of examination-induced stress on the technical performance of junior surgery residents.

METHODS: Twelve first-year surgery residents completed 2 surgical tasks (skin excision, tracheotomy) in low- and high-stress condition (in-training examination—Objective Structured Assessment of Technical Skills [OSATS]). Residents rated their subjective stress levels on a 10-point Likert-like scale. Performances were videotaped and assessed by 3 blinded experts using checklist and global rating scales.

RESULTS: Residents reported moderately higher stress levels in the exam condition than in the low-stress conditions ($P < .05$). Their performance was rated higher in the exam condition on the checklist scales ($P < .05$) but not on the global rating scales ($P = .79$).

CONCLUSIONS: Residency in-training exams induce moderate stress levels in junior surgery residents and are accompanied by improvements in technical performance as assessed by checklist-based scales. There were no differences on the global rating scales due to stress conditions, suggesting that residents were better at following the itemized sequence of movements when stressed, but their overall global performance was not altered.

© 2008 Elsevier Inc. All rights reserved.

Clinical practice can be filled with stress for physicians, resulting from treating critically ill and rapidly deteriorating patients, functioning while deprived of sleep, inadequate personal time, and disruptions in social support.^{1–4} During training, medical students and residents encounter addi-

tional stressors such as licensing exams and conflicts between their clinical and education responsibilities.⁵ Researchers have shown that the stress encountered during clinical practice and medical education has a significant effect on individuals, leading to higher than average levels of anxiety and depression in medical trainees compared to the general population of working adults.^{4–6} Despite the documented link between job stress and mental health,⁷ there is little research on the impact of stress on the cognitive, behavioral, and technical skills of health professionals.⁸

Friedlish et al⁹ recently showed that the inclusion of a laboratory-based technical skills station on a Canadian high-

This study was funded by a Royal College of Physicians and Surgeons of Canada medical education research grant. S.I. Woodrow was supported by a Royal College of Physicians and Surgeons of Canada Fellowship in Medical Education.

* Corresponding author. Tel.: +1-416-340-3054; fax: +1-416-340-3792.

E-mail address: vicki.leblanc@utoronto.ca

Manuscript received Apr 10, 2007; revised manuscript July 30, 2007

stakes examination for general license is both feasible and valid. To date we know very little about the technical performance of residents during this type of examination, as compared to performance on similar stations when not being evaluated. However, evaluations—whether summative or formative—are reported as being one of the most significant sources of stress for residents.² In light of the possibility of future inclusion of technical skills evaluation on high-stakes examination, it is important to understand how stress affects trainees' abilities to carry out necessary technical skills. Therefore, before the widespread implementation of technical skills assessment during a formal examination, it is crucial to study the possible changes in technical performance as a consequence of the stress and anxiety that can be imposed by these examinations.

The few studies in which the impact of stress on clinical performance has been investigated have focused on cognitive performance. Cumming and Harris¹⁰ found that both experimentally induced and naturalistic anxiety impaired the clinical decision-making ability of senior medical radiation science students. In a study with highly skilled paramedics, the stress induced by working through realistic clinical scenarios in a human patient simulator led to a significant decrease in the ability to solve drug calculation problems.¹¹ Both of these studies are consistent with the psychological literature indicating that socio-evaluative stressors, where one is being evaluated or critiqued by one or several other persons, reliably induce subjective stress responses in individuals.^{12,13} Interestingly, there are no documented studies evaluating the effects of stress on the motor skills of surgical trainees. In the motor performance literature, the findings regarding the effects of stress are ambiguous. While some researchers have found that induced stress impairs motor performance in sporting activities, other researchers have found that stress can facilitate motor performance.¹⁴

The purpose of the current study was to measure the effects of examination stress on the technical skills of surgery residents. The subjective stress levels and the technical performance of junior residents were measured during low-stress conditions and during in-training examinations.

Methods

Participants

The participants in this study were 12 surgery residents enrolled in their first year of postgraduate training (PGY1) at the University of Toronto. The residents participated in the study on a voluntary basis, in conjunction with their end of year in-training examination in a laboratory-based course of principles of surgery. Informed consent was obtained from each participant, and institutional ethics approval was obtained. The residents received a modest monetary honorarium for their participation in the study.

Study design

The study design was a within-subject repeated-measures design, with the residents serving as their own controls. Each resident performed 2 experimental tasks during their in-training Objective Structured Assessment of Technical Skills (OSATS) examination, and during 2 control low-stress conditions. The OSATS examination is considered a high-stress condition because it contains 2 elements of socio-evaluative stressors¹³; the presence of others for purposes of evaluation, and elevated stakes because students with extremely low scores are reported to the program director with recommendation for additional practice. Failure on this exam is documented in the resident's training portfolio and the resident undergoes remediation.

The first low-stress condition (pre-exam) occurred 1 week prior to the OSATS exam and the second low-stress condition (post-exam) occurred 1 week following the OSATS exam. We included 2 low-stress conditions, 1 before and 1 after the exam, in order to measure the effects of familiarity with the exam on performance, given that it was not feasible to counterbalance the order of the OSATS and low-stress conditions.

In the 2 low-stress conditions, the residents performed 4 tasks: 2 experimental procedures and 2 filler procedures. The filler procedures were included to ensure that the exact content of the OSATS examination was not revealed to the residents. During the pre-exam low-stress session, the residents were simply informed that any or all of the procedures could be included on their OSATS examination. Consistent with the format of the OSATS examination, and to avoid any potential confounds of time pressure, the residents were given 12 minutes to complete each task. Their performance of each task was videotaped from the elbows down, and the residents wore surgical gowns and gloves to achieve blinding of the resident identity and experimental condition. The same 4 procedures were performed during the post-exam low-stress session in order to keep the low-stress sessions as similar as possible.

During the OSATS examination, the residents participated in their regularly scheduled in-training exam in a standard manner. The OSATS consisted of an 8-station, bell-ringer format examination. During each station, residents were given 12 minutes to complete a defined surgical task on a bench model. Examiners (faculty surgeons, all fellows of the Royal College of Physicians and Surgeons of Canada and experienced with the OSATS examinations) rated the residents' performance on detailed task-specific checklists and global rating scales. These examiners rated the residents by direct observation for the purposes of the Department of Surgery OSATS examination, and these scores were not available for the purposes of this study. Performances on each of the selected tasks were videotaped from the elbows down. The residents wore surgical gowns and gloves and were identified by a unique code to achieve blinding of the resident identity and experimental condition.

Permission to videotape the performance of the residents on the selected tasks during this OSATS examination was granted by the University's research ethics board, the chair of the Department of Surgery, and the course director. As described in the "outcome measures" section, a separate group of examiners scored the videotapes for the purposes of this study.

Tasks

The residents performed 2 experimental tasks in each of the experimental conditions. The first task was a skin excision. The residents were required to excise a skin lesion with an elliptical incision and to close the wound primarily with interrupted sutures. The second task was a tracheotomy. The residents were required to perform all major steps of a tracheotomy including (1) skin incision; (2) dissection through strap muscles; (3) identification of trachea; (4) tracheotomy; (5) coordination of endotracheal tube withdrawal with proper tracheotomy tube insertion; and (6) appropriately securing tracheotomy tube. The 2 particular tasks were selected because they were part of the first-year curriculum of the surgical residency training and the residents had received approximately 6 hours of laboratory-based training and practice on these tasks. Furthermore, we selected tasks that were reported as having different levels of difficulty, with the tracheotomy reported as being more complex and challenging. While the residents were familiar and experienced with the tasks, they were not informed of the items included in the performance measures.

Outcome measures

The subjective stress levels were assessed by a question posed to the residents prior to each of the procedures and a question posed immediately after each procedure. The question posed prior to each procedure was: "How stressed do you feel right now?" The question posed immediately following each procedure was: "How stressful was the task you just completed?" The residents answered each question by circling a number on an anchored scale of 1 to 10 (1 = not at all, 10 = extremely). These short questions have been shown to be sensitive measures of subjective responses to various stressors,^{15,16} and to be comparable to multi-item subjective assessments and more invasive measures of physiological stress such as cortisol and alpha-amylase levels.¹⁷

The videotaped performances were randomized and then independently viewed and rated by three raters (surgeons who had a minimum of 1 year experience with rating technical performance) using both checklist scales and global ratings of performance.¹⁸⁻²¹ The research raters had similar experience with the checklist and global ratings scales as did the examiners for the OSATS examination. Each resident was awarded a checklist score, and a global rating score. The checklist score indicates the completion of a

detailed list of operation specific procedures. It consists of a checklist of the chronological steps involved in a procedure. For each step, residents are scored a 0 if the step is incorrectly performed and a 1 if it is accurately performed.¹⁸ The residents are awarded a score between 0 and 100, depending on the percentage of steps properly performed during each task. The global rating score is a measure of general operative performance. It consists of seven items (ie, respect for tissue) that are scored on a 5-point scale. Points 1, 3, and 5 are anchored with behavioral markers.¹⁸ Residents are awarded a score between 7 and 35, based on the sum of the scores on each item. Both of these methods of evaluation have been previously validated in numerous studies.¹⁹⁻²¹ The raters were blinded to the experimental condition and to the residents' identities.

Analyses

Each of the subjective stress measures and performance measures were submitted to separate repeated-measures analyses of variance (ANOVA), with the procedure (skin excision, tracheotomy) as a random effect repeated measure and condition (pre-exam low stress, OSATS, post-exam low stress) as a fixed effect repeated measures. Post hoc analyses were conducted with the use of repeated measures *t* tests. Agreement between the raters on the checklist and global rating scales was computed by means of intraclass correlation coefficients. The data are presented as the means and standard errors, unless otherwise noted.

Results

Stress responses

Across the 3 conditions, the residents reported higher subjective stress levels prior to performing the tracheotomy than prior to performing the skin excision. For both tasks, the residents reported higher subjective stress levels prior to performing the OSATS exam than prior to the 2 low stress conditions. They also reported slightly higher stress levels prior to the pre-exam low-stress session than prior to the post-exam stress session, which likely reflected their uncertainty or lack of experience with the set up of the study. There was no significant task by condition interaction. **Table 1** provided means and standard errors, as well as a summary of the statistical analyses.

Following completion of the procedures, the residents rated the tasks as having been more stressful in the OSATS condition than in the 2 low-stress conditions. The residents also rated the tasks as having been less stressful in the post-exam than in the pre-stress low condition. They reported the 2 tasks as having been equally stressful. There was no significant task by condition interaction. **Table 2** lists the means and standard errors, and summarizes the statistical analyses.

Table 1 Subjective stress question 1: how stressed do you feel right now?

	Pre-exam low stress	OSATS high stress	Post-exam low stress	Average
Suture	1.9 (.4)	3.3 (.5)	1.3 (.2)	2.2 (.2)
Tracheotomy	2.4 (.3)	4.1 (.6)	1.7 (.2)	2.7 (.2)
Average	2.2 (.3)	3.7 (.5)	1.5 (.2)	

Scores are out of 10. Numbers in parentheses are SEM, as generated by ANOVA.

Main effect of task: $F(1,11) = 4.95$, $MSE = 1.0$, $P < .05$, $\eta^2 = .31$. Tracheotomy > suture.

Main effect of condition: $F(2,22) = 12.0$, standard error of the mean = 2.5, $P < .01$, $\eta^2 = .52$. OSATS > pre low stress, post low stress; pre low stress > post low stress.

No interaction between task and condition: $F(2,22) = 2.9$, $MSE = .9$, $P = .75$, $\eta^2 = .03$.

Performance

The agreement between the raters on the checklist scales was good, with the inter-rater reliability (intraclass correlation coefficient) of .70.²² The scores from the 3 raters were averaged together and all analyses were conducted on these averaged scores. The analyses of the checklist scores revealed a significant main effect of condition. The scores during the OSATS examination were higher than in the pre-exam low-stress condition. The scores were not statistically different between the 2 low-stress conditions or between the OSATS examination and the post low-stress conditions. Checklist scores on the suture and the tracheotomy tasks did not differ significantly from each other. There was no significant condition by task interaction on the checklist scores. Table 3 provides means and standard errors, as well as a summary of the statistical analyses.

The agreement between the raters on the global rating scales was fair, with the inter-rater reliability (intraclass correlation coefficient) of .52.²² The scores from the 3 raters

Table 2 Subjective stress question 2: how stressful was the task you just completed?

	Pre-exam low stress	OSATS high stress	Post-exam low stress	Average
Suture	2.3 (.4)	2.6 (.6)	1.6 (.2)	2.5 (.3)
Tracheotomy	3.0 (.6)	4.8 (.6)	1.3 (.1)	3.1 (.3)
Average	2.7 (.4)	4.2 (.6)	1.4 (.1)	

Scores are out of 10. Numbers in parentheses are SEM, as generated by ANOVA.

No main effect of task: $F(1,11) = 3.81$, $MSE = 1.5$, $P = .08$, $\eta^2 = .25$.

Main effect of condition: $F(2,22) = 16.8$, $MSE = 2.7$, $P < .01$, $\eta^2 = .61$. OSATS > pre low stress, post low stress; pre low stress > post low stress.

No interaction between task and condition: $F(2,22) = 2.8$, $MSE = 1.2$, $P = .08$, $\eta^2 = .21$.

Table 3 Average scores on the checklist scale

	Pre-exam low stress	OSATS high stress	Post-exam low stress	Average
Suture	78.9 (2.4)	8.9 (2.5)	81.7 (1.9)	8.5 (1.9)
Tracheotomy	77.8 (2.5)	83.9 (2.6)	79.0 (3.3)	8.2 (2.4)
Average	78.3 (1.7)	82.4 (2.2)	8.3 (1.9)	

Scores are out of 100%. Numbers in parentheses are SEM, as generated by ANOVA.

No main effect of task: $F(1,11) = .1$, $MSE = .1$, $P = .91$, $\eta^2 = .001$.

Main effect of condition: $F(2,22) = 3.4$, $MSE = .003$, $P = .05$, $\eta^2 = .23$. OSATS > pre low stress.

No interaction between task and condition: $F(2,22) = 1.3$, $MSE = .004$, $P = .29$, $\eta^2 = .11$.

were averaged together and all analyses were conducted on these averaged scores. The analyses of the global rating scores revealed a main effect of condition. The residents' scores did not differ between the OSATS examination and the pre-exam low-stress condition. The scores during the post-exam low-stress condition were higher than the scores in both the pre-exam low-stress condition and the OSATS examination, suggesting a slight improvement due to practice (ie, exposure to the experimental examinations). Global rating scores for the suture and tracheotomy task did not differ significantly from each other. There was no significant task by condition interaction. Table 4 lists the means and standard errors, and summarizes the statistical analyses.

Comments

In this study, the subjective stress levels and the technical performance of junior residents were measured during low-stress conditions and during in-training examinations. The effects of examination-induced stress on the technical skills of surgery residents were assessed. The results showed that a laboratory-based in-training examination led to moderate increases in the subjective stress levels of junior surgery

Table 4 Average scores on the global rating scale

	Pre-exam low stress	OSATS high stress	Post-exam low stress	Average
Suture	25.6 (.6)	24.6 (1.2)	27.0 (.9)	25.7 (.7)
Tracheotomy	24.4 (1.2)	25.9 (1.3)	28.4 (.8)	26.2 (.9)
Average	25.0 (.7)	25.2 (1.0)	27.7 (.6)	

Scores are out of 35. Numbers in parentheses are SEM, as generated by ANOVA.

No main effect of task: $F(1,11) = .2$, $MSE = 21.7$, $P = .67$, $\eta^2 = .02$.

Main effect of condition: $F(2,22) = 5.7$, $MSE = 9.23$, $P < .01$, $\eta^2 = .34$. Post low stress > OSATS, pre low stress.

No interaction between task and condition: $F(2,22) = 2.3$, $MSE = 5.8$, $P = .12$, $\eta^2 = .18$.

residents. This socio-evaluative stress induced by the in-training examinations appeared to have a facilitating effect on certain aspects of the technical performance. Specifically, we observed an improvement of checklist scores during the examination compared to the low stress pre-exam session. These checklist scores represent a measure of whether the appropriate actions were taken and accurately performed during the procedure. These findings cannot be attributed to repeated examinations effects, because performance during the post-exam low-stress condition was not significantly different than performance during the pre-exam low-stress condition.

In contrast, the residents' efficiency, economy, and fluidity of motion, as assessed by the global rating scale, do not appear to be influenced by moderate stress levels. Rather, the scores of the residents were higher in the post-exam low-stress condition than in both the pre-exam low-stress condition and the in-training examination. Given that the residents reported that the tasks as more stressful in the pre-exam low-stress and the in-training examination condition than in the post-exam low-stress condition, it is possible that the more global aspects of surgical performance may be vulnerable to slight increases in stress levels. However, if increased stress levels led to impaired performance, we would have expected the performance decrements to be substantially greater in the exam condition, in which the residents reported significantly higher levels of subjective stress than in the pre-exam low-stress session, both in anticipation of the task and following the task. A more likely explanation is that global performance improved with the repeated execution of the procedures. This aspect of performance may be governed by different parts of the brain than the knowledge of appropriate actions. It is considered procedural knowledge—the "know-how" to perform tasks.^{23,24} This knowledge can generally not be verbally articulated and is nonconscious. Consistent with our results, previous research has demonstrated that this procedural knowledge is improved with repeated practice.^{23,24} In summary, junior residents appear to be better at remembering and following the sequenced steps of a procedure (declarative knowledge) under moderate stress, but they do not appear to be better in executing and performing these steps (procedural knowledge). Rather, procedural knowledge is improved with repetition and practice.

Interestingly, the 2 selected tasks did not differ in terms of how stressful the residents had found them or in terms of the performance scores (checklist and global rating scores). The anecdotal reports of residents and educators suggested that the tracheotomy would be more complex than the suture task. The participating residents echoed this belief by reporting higher anticipatory stress levels prior to the tracheotomy than prior to the suture task. However, the results of the study suggest that the 2 tasks, as assessed in the laboratory-based skills courses, are equal in terms of how residents perform on them and in terms of how stressful they are to perform.

The findings from this study are consistent with the literature on the effects of stress on motor performance from

other domains. While it is generally accepted that stress, both acute and chronic, has a negative effect on performance, there is accumulating evidence that stress can sometimes facilitate performance. The effects of socio-evaluative stressors appear to be mediated by the level of difficulty of a task.^{25,26} The presence of observers or examiners is thought to influence behavior in a number of ways. Theories of general drive propose that the presence of others increases the general arousal of the individual performing a task.²⁵ In situations of increased arousal, individuals are more likely to engage in behavior that is well-learned or automatic. On easy tasks, in which the behavior to be performed is over-learned and often automatized, this will lead to an improvement in performance. On tasks that are more complex, the appropriate behavior is not well-learned and some incorrect actions are present. Increased arousal increases the occurrence of these incorrect or oversimplified actions along, thereby impairing performance.

The presence of others can also influence the cognitions of the individual performing a task. Self-attention theories²⁶ suggest that the presence of others causes the individual to become more aware of his or her own behavior and of the salient standard of behavior (the "ideal"). This causes the individual to engage in a self-regulating process by which he or she is more focused on attempting to match the present behavior with the standards of behavior. If the task is an easy one, the discrepancy between the current behavior and the standard will be small. In such a case, the individual develops a positive expectation of outcome and performance is enhanced as the individual attempts to improve performance to meet the standard. However, if the task is difficult or complex, the individual is more likely to perceive a large discrepancy between the current behavior and the standard of performance. In such a case, the individual develops a negative expectation of outcome for resolving the discrepancy and is more likely to withdraw from the task, leading to performance impairment.

Regardless of the mechanisms by which behavior is affected (increased arousal, increased self-attention), the findings consistently suggest that the presence of others facilitates performance on easy tasks and impair it on difficult tasks.¹³ The specific actions required on the tasks selected for the current study appeared relatively easy for the residents, as evidenced by an average score (for both tasks) on the checklist scales of performance of around 80%. Thus, our findings are consistent with the literature of socio-evaluative stressors.

The results of this study add to the literature by clarifying the aspects of performance that are improved with moderate stress levels. The improvements on the checklist scale suggest that it is the residents' ability to remember and follow the sequenced steps of procedures that is enhanced by moderate stress. In other words, their declarative memory^{23,24}—the knowledge of facts—is improved. This is consistent with previous research demonstrating that moderate stress levels lead to increases in cortisol, a hormone that is heavily

involved in the declarative memory processes in the brain.²⁷ In turn, this can lead to improvements in declarative memory.²⁸ In contrast, the more global aspects of performing a skills, the residents' fluidity, economy, and efficiency of motions, does not appear to be influenced by moderate stress levels. Rather, they appear to improve with practice.

The results of this study have interesting implications for surgical education. Surgeries on live patients can be stressful for junior residents. However, the results suggest that if residents are well-trained on the procedures that they must complete during these surgeries, their performance may be immune to the stress they experience during these surgeries. This provides support for the importance of laboratory-based surgical training in preparation for operating room-based surgical training. Evidently, these conjectures need to be taken with caution, as the study was conducted with junior residents who were completing relatively simple tasks under moderately stressful conditions. Further studies need to be conducted with more senior residents, with tasks of varying levels of difficulty or mastery by the resident, and under various conditions and levels of stress.

Furthermore, the recent interest in incorporating technical skill stations into high-stakes examinations suggest a second area of inquiry that merits further research. The stress levels induced by in-training examinations appear to be moderate. Those induced by high-stakes licensing examinations are likely to be higher, due to the implications associated with the level of performance on such examinations. There is some evidence suggesting that the more evaluative an audience, the greater the performance impairment.²⁹ Given the interaction between stress and the degree of difficulty of the task being performed, it is possible that candidates of varying levels of competence may be differently affected by the stress of high-stakes examinations. Those trainees who have mastered a particular task may perform better during the exams while those who have not mastered the task may demonstrate impairments due to the stress of the examination, thus exaggerating the differences in the final marks awarded to these 2 types of trainees.¹³

In summary, the findings from this study suggest that in-training examinations induce moderate stress in junior residents. These stress levels are accompanied by improvements in following the itemized sequence of movements during technical procedures. To fully understand the effects of socio-evaluative stressors on technical performance, further research should be aimed at looking at the contributions of the complexity of the tasks, the stakes of an examination, and the level of experience or expertise of the individuals.

References

1. Butterfield PS. The stress of residency: a review of the literature. *Arch Intern Med* 1988;148:1428–35.
2. McCue JD, Sachs CL. A stress management workshop improves residents' coping skills. *Arch Intern Med* 1991;151:2273–7.
3. Michie S, Williams S. Reducing work related psychological ill health and sickness absence: a systematic literature review. *Occup Environ Med* 2003;60:3–9.
4. Reuben DB. Depressive symptoms in medical house officers: effects of level of training and work rotation. *Arch Intern Med* 1985;145:286–8.
5. Toews JA, Lockyer JM, Dobson DJG, et al. Stress among residents, medical students, and graduate science (MSc/PhD) students. *Acad Med* 1993;68(suppl):S46–8.
6. Peterlini M, Tiberio IFLC, Saadeh A, et al. Anxiety and depression in the first year of medical residency training. *Med Educ* 2002;36:66–72.
7. French JRR Jr, Caplan RD, Van Harrison R. *The Mechanisms of Job Stress and Strain*. Chichester, NY: John Wiley & Sons; 1982.
8. Schull MJ, Ferris LE, Tu JV, et al. Problems for clinical judgment: 3. Thinking clearly in an emergency. *Can Med Assoc J* 2001;164:1170–5.
9. Friedlish M, Wood T, Smee S. Is it feasible to include a technical skills station on a national licensing examination? *Association for Surgical Education Proceedings*, 2006:53.
10. Cumming SR, Harris LM. The impact of anxiety on the accuracy of diagnostic decision-making. *Stress Health* 2001;17:281–6.
11. LeBlanc VR, McArthur B, King K, et al. Stressful clinical scenarios impair paramedics' ability to perform drug calculations. *Prehospital Emergency Care* 2004;8:83.
12. Kemeny ME. The psychobiology of stress. *Curr Direct Psychol Sci* 2003;12:124–9.
13. Driskell JE, Salas E. *Stress and Human Performance*. Mahwah, NJ: Lawrence Erlbaum; 1996.
14. Neiss R. Reconceptualizing arousal: psychobiological states in motor performance. *Psychol Bull* 1988;103:345–66.
15. Tomaka J, Blascovich J, Kelsey RM, et al. Subjective, physiological, and behavioral effects of threat and challenge appraisal. *J Pers Soc Psychol* 1993;65:248–60.
16. Tomaka J, Blascovich J, Kibler J, et al. Cognitive and physiological antecedents of threat and challenge appraisal. *J Pers Soc Psychol* 1997;73:63–72.
17. Chatterton RT, Vogelsong KM, Lu Y-C, et al. Hormonal responses to psychological stress in men preparing for skydiving. *J Clin Endocrinol Metab* 1977;82:2503–9.
18. Reznick R, Regehr G, MacRae H, et al. Testing technical skill via innovating "bench station" examination. *Am J Surg* 1997;173:226–30.
19. Martin JA, Regehr G, Reznick R, et al. Objective Structured Assessment of Technical Skills (OSATS) for surgical residents. *Br J Surg* 1997;173:226–30.
20. Faulkner H, Regehr G, Martin J, et al. Validation of an objective structure assessment of technical skill for surgical residents. *Acad Med* 1996;71:1363–5.
21. Backstein D, Agnidis Z, Sidhu R, et al. Effectiveness of repeated video feedback in the acquisition of a surgical technical skills. *Can J Surg* 2005;48:195–200.
22. Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd Ed. New York: Wiley; 1981.
23. Willingham DT. *Cognition: The Thinking Animal*. 2nd Ed. Upper Saddle River, NJ: Pearson Prentice Hall; 2004.
24. Reed SK. *Cognition: Theory and Application*. 6th Ed. Belmont, CA: Thomson Learning. 2004.
25. Cottrell NB. Social facilitation. In McClintock CG, editor. *Experimental Social Psychology*. New York: Holt; 1972;185–236.
26. Carver CS, Scheier MF. Self-attention induced feedback loop and social facilitation. *J Exp Soc Psychol* 1981;17:545–68.
27. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol Bull* 2004;130:355–91.
28. Erickson K, Drevets W, Schulin J. Glucocorticoid regulation of diverse cognitive functions in normal and pathological states. *Neurosci Biobehav Rev* 2003;27:233–46.
29. Paulus PB, Murdoch P. Anticipated evaluation and audience presence in the enhancement of dominant responses. *J Exp Soc Psychol* 1971;7:280–91.