World Journal of Surgery

Development and Validation of a Surgical Workload Measure: The Surgery Task Load Index (SURG-TLX)

Mark R. Wilson · Jamie M. Poolton ·
Neha Malhotra · Karen Ngo · Elizabeth Bright ·
Rich S. W. Masters

Abstract

Background The purpose of the present study was to develop and validate a multidimensional, surgery-specific workload measure (the SURG-TLX), and to determine its utility in providing diagnostic information about the impact of various sources of stress on the perceived demands of trained surgical operators. As a wide range of stressors have been identified for surgeons in the operating room, the current approach of considering stress as a unidimensional construct may not only limit the degree to which underlying mechanisms may be understood but also the degree to which training interventions may be successfully matched to particular sources of stress.

Methods The dimensions of the SURG-TLX were based on two current multidimensional workload measures and developed via focus group discussion. The six dimensions were defined as mental demands, physical demands, temporal demands, task complexity, situational stress, and distractions. Thirty novices were trained on the Fundamentals of Laparoscopic Surgery (FLS) peg transfer task and then completed the task under various conditions designed to manipulate the degree and source of stress experienced: task novelty, physical fatigue, time pressure, evaluation apprehension, multitasking, and distraction.

Results The results were supportive of the discriminant sensitivity of the SURG-TLX to different sources of stress. The sub-factors loaded on the relevant stressors as hypothesized, although the evaluation pressure manipulation was not strong enough to cause a significant rise in situational stress.

Conclusions The present study provides support for the validity of the SURG-TLX instrument and also highlights the importance of considering *how* different stressors may load surgeons. Implications for categorizing the difficulty of certain procedures, the implementation of new technology in the operating room (man–machine interface issues), and the targeting of stress training strategies to the sources of demand are discussed. Modifications to the scale to enhance clinical utility are also suggested.

M. R. Wilson · J. M. Poolton · N. Malhotra · K. Ngo ·
R. S. W. Masters
Institute of Human Performance, University of Hong Kong,
Hong Kong, SAR, People's Republic of China

M. R. Wilson (✉)
School of Sport and Health Sciences, University of Exeter,
St. Luke's Campus, Exeter EX1 2LU, United Kingdom
e-mail: mark.wilson@ex.ac.uk

J. M. Poolton
Department of Surgery, University of Hong Kong, Hong Kong,
SAR, People's Republic of China

E. Bright
Department of Urology, Royal Devon and Exeter Hospital,
Royal Devon and Exeter NHS Foundation Trust, Barrack Road,
Exeter EX2 5DW, United Kingdom

Introduction

The surgical operating room is a multifaceted environment that exposes operating surgeons and their teams to considerable stress-inducing conditions. Challenges, such as procedure complexity, time pressure, peer evaluation, multitasking, and distractions all have the potential to raise levels of intraoperative stress [1, 2]. Despite the multiple stressors that surgeons may face, they are more likely to deny potential effects of stress on their performance than individuals in other challenging environments [3]. Such an attitude has discouraged applied research in the field and limited organizational and educational change policies [4].

As intraoperative stressors are seldom factored in as potential contributors to surgical outcome, there are also significant negative implications for patient care and safety.

Stress is experienced when perceived resources are outweighed by demands [5, 6]. Given that multiple sources of stress have been identified, one weakness of current research is that it adopts a unidimensional approach to measurement. While validated instruments such as the State Trait Anxiety Inventory (STAI) [7] provide a measure of emotion (anxiety), other mechanisms may underpin the stress-performance relationship. Indeed, different stressors are likely to cause surgical performance to break down for different reasons. Considering stress as a unidimensional construct not only limits the degree to which underlying mechanisms may be understood but also the degree to which a training intervention may be successfully matched to a particular source of stress.

Few studies in surgery have attempted to gain insight into the specific demands imposed on surgeons by typically experienced stressors. In the fields of aviation and industrial ergonomics, however, the study of mental demand (workload) has been a major area of inquiry, as researchers have sought to examine the potential causes of poor performance linked to increased workload [8–11]. Workload is a multifaceted construct, determined by the interaction of the task demands, the circumstances under which the task is performed, and the skills, behaviors, and perceptions of the individual [12, 13]. It is apparent from this definition that anxiety (stress) may be but one factor with an impact on the demands of the task.

The most widely used measure of workload in human factors research has been the NASA-Task Load Index (NASA-TLX) [14], is a multidimensional rating scale that has six bipolar dimensions: mental demand (MD); physical demand (PD); temporal demand (TD); own performance (P); effort (E); and frustration (F). The dimensions therefore reflect task-related (MD, PD, TD), subject-related (P), and behavior-related (F and E) factors. While multidimensional measures provide stronger diagnosticity (i.e., the capability of an instrument to discriminate between different *types* of workload [9, 13]), a weakness is that they are generally created for a specific environment or task, and therefore may not reflect different dimensions of workload in other environments [15].

Although the NASA-TLX has been adopted as a measure of workload in recent surgical research [16–20], in all cases the individual dimension scores were simply aggregated to provide a total workload measure. This process ignores the primary advantage of multidimensional scales: their ability to discriminate between different sources of workload. The purpose of the present study was therefore to develop and validate a surgery-specific version of the Task Load Index (SURG-TLX), and to determine whether it provides diagnostic information regarding the impact of various sources of stress on the perceived demands of trained surgical operators.

Methods

Scale development

As the NASA-TLX is a well-validated instrument [21, 22], the intention was to maintain its general structure but make it more relevant to the specific demands of surgery [15]. The first step was to consider the process adopted in developing another TLX variant, designed for car driving; the Driving Activity Load Index (DALI) [23]. The DALI's six dimensions (effort of attention, visual demand, auditory demand, temporal demand, interference, and situational stress) were first determined by discussion with a number of experts in driving research. A study was then designed to test the sensitivity and diagnosticity of the instrument for typical driving tasks; interacting with a navigation system and operating a hands-free car phone. Results confirmed that the DALI dimensions were sensitive to these manipulations [23].

To develop a surgery-specific version of the NASA-TLX, we consulted qualitative research that has identified key intraoperative stressors [2] and considered which dimensions of the NASA-TLX and DALI best approximate the demands faced by surgical operators. The three task demand dimensions from the NASA-TLX were retained (mental, physical, and temporal demands), as were the environmental demand dimensions from the DALI (distractions and situational stress). It was felt that a final dimension reflecting Task Complexity was more appropriate than one related to effort or frustration. The specific dimensions for the SURG-TLX were therefore formulated and defined as follows:

1. *Mental demands*: How mentally fatiguing was the procedure?
2. *Physical demands*: How physically fatiguing was the procedure?
3. *Temporal demands*: How hurried or rushed was the pace of the procedure?
4. *Task complexity*: How complex was the procedure?
5. *Situational stress*: How anxious did you feel while performing the procedure?
6. *Distractions*: How distracting was the operating environment?

Eight experienced surgeons from a range of disciplines (four Consultants and four Specialist Registrars) were asked to provide their opinions of the SURG-TLX's

dimensions, as well as provide "free" comments about which factors made procedures demanding. While a variety of specific factors were raised (e.g., negativity from others in the operating room, nonavailability of preferred equipment, patient expectations) there was general agreement that the dimensions were reflective of the typical demands experienced in surgery. The surgeons were provided with the NASA-TLX and DALI dimensions for comparison, and all 8 agreed that mental demands, temporal demands, task complexity, and distractions were important factors affecting workload judgments. Two of the Consultants felt that physical demands and situational stress may not be as relevant to workload as the frustration dimension from the NASA-TLX. However, because most of the surgeons were satisfied with the dimensions selected, we decided to maintain the original six-dimension structure of the index.

Having developed the instrument, the second phase of the study aimed to validate it by exposing trainee operators to various intraoperative stressors as they performed a well-validated laparoscopic task.

Subjects

Novices ($n = 30$ medical students) volunteered to take part in the research. Institutional ethical approval was obtained prior to the commencement of the study, and all subjects provided written informed consent and demographic information before testing. Subjects were informed that they would be given the opportunity to perform a laparoscopic task under a variety of conditions in a laboratory supporting clinical simulation. Subjects attended individually and were paid $HK150 for taking part.

Materials and task

The task adopted was the validated Fundamentals of Laparoscopic Surgery (FLS) peg transfer task [24]. The FLS training program model is endorsed by the American College of Surgeons and the Society of American Gastrointestinal and Endoscopic Surgeons, and consists of five tasks of increasing complexity [25, 26]. In the peg transfer task, six plastic objects are grasped, transferred, and positioned on a pegboard. Specifically, each object is picked up with grasper forceps from a pegboard on the surgeon's left, transferred in space to a grasper in the right hand and then placed around a post on the right-hand side of the pegboard. After all six objects have been transferred from left to right the process is reversed, requiring transfer from the right hand to the left hand. The exercise is timed and a penalty score is assessed whenever an object is dropped outside the surgeon's view.

As with the original TLX and the DALI, a two-part evaluation is required to complete the SURG-TLX. The first part involves calculating weights of the six dimensions following a set of 15 paired comparisons. The dimension with the highest weight is the most important contributing factor for the perceived workload (scores range from 0 to 5). The second part involves rating six bipolar scales reflecting the separate dimensions on a 20-point Likert scale, anchored between low and high (see Appendix 1 for the SURG-TLX). A workload score for each dimension is then calculated by determining the product of these two numbers. For example, a weight score of 4 and a rating of 15 equate to a workload score of 60 (scores range from 0 to 100). A total workload score is also determined by aggregating the scores from the six dimensions.

Procedure

Before training commenced, subjects watched an introductory video showing an expert complete the peg transfer task. They were then required to perform repetitions of the task until they reached proficiency; defined as completing the task in less than 54 s and without a penalty score on two consecutive trials and on ten additional nonconsecutive trials. Developers of the FLS skills curriculum [26] have recommended that surgical educators adopt this criterion for task proficiency, which is based on expert levels of performance [25]. Subjects were informed of the proficiency requirements at the outset of training and were offered feedback on their completion times whenever it was asked for.

The procedure consisted of training and testing phases. In the training phase, subjects trained on the peg transfer task for up to 90 min, or until proficiency was reached. Subjects completed the SURG-TLX after their fifth learning attempt (task novelty condition) and were asked to complete it with respect to their previous two attempts. Subjects also completed the SURG-TLX after their final attempt of this training session (physical fatigue condition). Again, subjects were asked to complete the instrument with respect to their previous two attempts. If proficiency was not attained in this time-frame then a second training session was organized for the following day. Sixteen of the 30 subjects had to complete an additional training session in order to reach the criterion level of proficiency. All subjects reached proficiency, taking on average 59.4 ($SD = 20.8$) trials to reach the criterion level of performance.

The testing phase was scheduled for the day after proficiency had been reached. Subjects first had to attain two consecutive criterion level completions. They then performed two trials in a control condition and each of three test conditions designed to simulate typical stressors

experienced during surgical performance [1] (counterbalanced design). The test conditions consisted of a multitasking condition, an evaluative condition, and a time pressure condition. The SURG-TLX was completed after the second trial of each condition in the test phase.

In the control condition subjects were simply asked to do their best in completing the task. The multitasking condition was designed to be distracting and mentally demanding, as subjects were required to perform mental arithmetic while completing the peg transfer task [4, 27–29]. Specifically, on the first trial subjects started counting back from 737 in sevens, and on the second trial they started counting from whichever number they reached on the first trial.

The evaluative condition involved a manipulation designed to increase ego-threat and performance anxiety [28, 30]. Subjects were informed that their performance was to be videotaped so it could be viewed by three of their course tutors and compared to the performance of trainee surgeons from the United Kingdom and the United States of America. The subjects were made aware of a video camera being turned on and were asked to say their name and year of study for the camera prior to completing their two trials. The final condition was designed to create an element of time pressure [4, 28]. Subjects were informed that some surgeries have to be completed under time constraints, perhaps because of complications occurring during the procedure. They were informed of their best time during training and were instructed to try to complete the task more quickly than on that attempt.

Data analysis

A mean workload score for each dimension (and total workload) was computed for each of the six conditions of interest (training phase: task novelty and physical fatigue; test phase: control, multitasking, evaluation and time pressure) and subjected to one way analysis of variance (ANOVA). Significant main effects were followed up with Bonferroni adjusted paired sample t-tests, and effect sizes were reported as partial eta squared (η_p^2).

Hypotheses

A series of hypotheses were developed based on the expected effects of the manipulations affecting workload (compared to the control condition):

Hypothesis 1: Primarily the Task Complexity dimension will be raised in the "task novelty" condition, reflecting the fact that the task is unfamiliar and unpracticed.

Hypothesis 2: Primarily the Physical Demands (fatigue) dimension will be raised in the "physical fatigue" condition, as subjects will have completed up to 90 min of training of a novel task.

Hypothesis 3: Primarily the mental demands and distraction dimensions will be raised in the multitasking condition, due to concurrent task loading.

Hypothesis 4: Primarily the situational stress dimension will be raised in the "evaluation" condition, due to the ego-threatening nature of the instructions.

Hypothesis 5: Primarily the temporal demands dimension will be raised in the "time pressure" condition.

Results

Task complexity

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 45.5$, $P < .001$, $\eta_p^2 = .61$. Bonferroni follow-up tests revealed that the multitasking condition was perceived to be significantly more complex than any other condition ($Ps < .001$; see Fig. 1). No other significant differences were evident.

Physical demands

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 13.0$, $P < .001$, $\eta_p^2 = .31$. Bonferroni follow-up tests revealed that physical demands were significantly higher in the physical fatigue condition than all other conditions (all $Ps < .05$). Furthermore, the multitasking condition was perceived as being significantly less physically demanding than all other conditions (all $Ps < .005$; see Fig. 1).
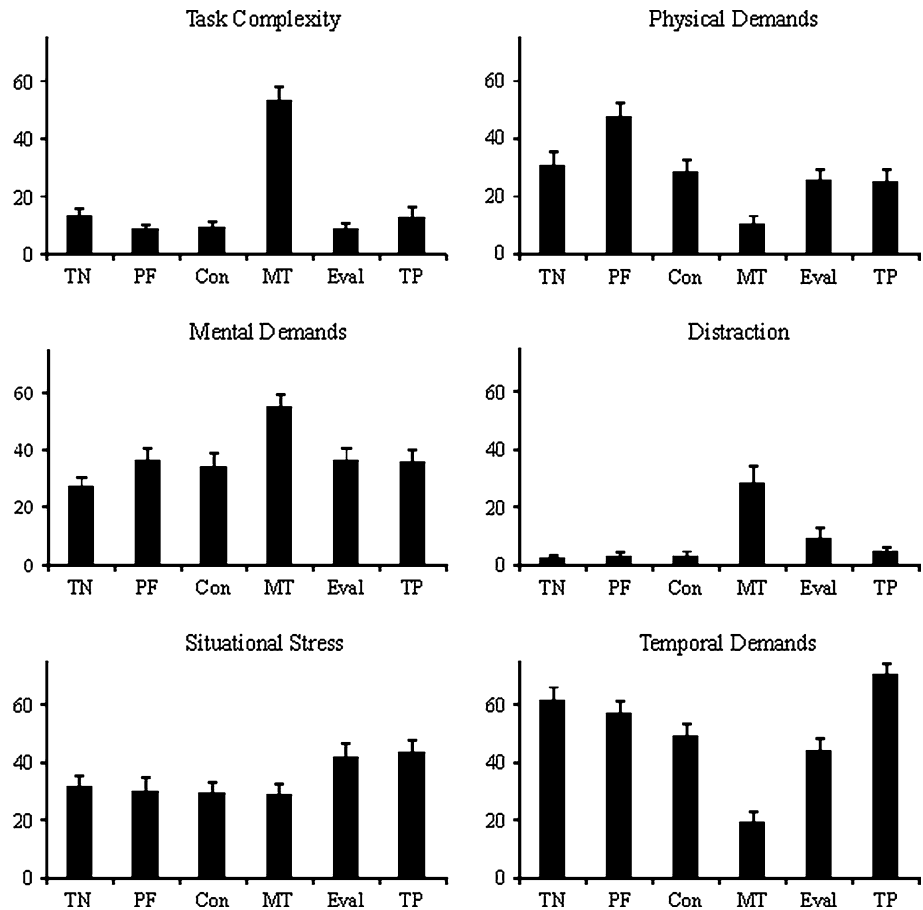
Mental demands

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 8.3$, $P < .001$, $\eta_p^2 = .22$. Bonferroni follow-up tests revealed that the reported mental demand in the multitasking condition was significantly higher than in all other conditions (all $Ps < .05$; see Fig. 1).

Distraction

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 12.7$, $P < .001$, $\eta_p^2 = .31$. Follow-up Bonferroni tests revealed that the multitasking condition was significantly more distracting than all other conditions (all $Ps < .05$; see Fig. 1). No other significant differences were evident.

Fig. 1 Mean workload score for each dimension of the SURG-TLX for the task novelty (*TN*), physical fatigue (*PF*), control (*Con*), multitasking (*MT*), evaluation (*Eval*), and time pressure (*TP*) conditions

Situational stress

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 3.1$, $P < .05$, $\eta_p^2 = .10$. Follow-up Bonferroni tests revealed that the time pressure condition was most stressful, although this was only at a significant level when compared to the multitasking condition ($P < .05$) and the control condition ($P < .05$, see Fig. 1).

Temporal demands

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 28.6$, $P < .001$, $\eta_p^2 = .50$. Bonferroni follow-up tests revealed that subjects perceived the time pressure condition to have significantly higher temporal demands than all other conditions (all $Ps < .05$), except the novel task condition ($P = .49$). The temporal demands of the multitasking condition were also perceived to be significantly less than all other conditions (all $Ps < .001$, see Fig. 1).

Total workload

Analysis of variance revealed a significant main effect for condition, $F (5,145) = 7.68$, $P < .001$, $\eta_p^2 = .21$.

Follow-up Bonferroni tests revealed that the control condition was significantly less demanding than "physical fatigue" ($P < .05$), multitasking ($P < .005$), and time pressure ($P < .001$) conditions. The time pressure condition was also significantly more demanding than the evaluation condition ($P < .001$; see Fig. 2).

Discussion

The aim of this research was to develop and validate a surgery-specific, multidimensional workload measure (the SURG-TLX), based on the NASA-TLX [14] developed for
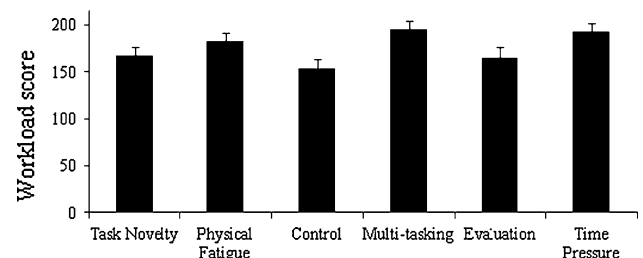

Fig. 2 Mean total workload scores for the six performance conditions of interest

pilots. The advantage of multidimensional measures is that they provide a degree of diagnosticity, although this is at the expense of specificity [13, 15]. While the original TLX has been adopted in surgical settings, only the total workload data have been presented [16–20], limiting the utility of the instrument to provide insights into *specific* sources of workload. Given that a wide range of stressors have been identified for surgeons in the operating room [1, 13] a surgery-specific workload measure might provide useful information to categorize procedures, guide training, and design stress management interventions.

The results of the present study, using recently trained laparoscopic operators and a validated laparoscopic task, revealed that the SURG-TLX is sensitive to a variety of different surgical stressors; including physical fatigue, time pressure, multitasking, and increased complexity. Indeed, of the five hypotheses developed to test the sensitivity of the six dimensions, there were only two somewhat unexpected, but explainable, results. We expected that, as relative novices (five trials of laparoscopic training), subjects would consider the task to be demanding [31] (high task complexity; hypothesis 1). However, only the multitasking condition was perceived to be significantly more complex than the control condition (Fig. 1a). Although this was not an a priori prediction, it is perhaps not surprising that subjects found the task to be more complex when a concurrent cognitive load was added.

The other unexpected finding was that the situational stress dimension (Fig. 1e) was not significantly higher in the evaluative condition (hypothesis 4). Previous research has demonstrated that trainee surgeons find evaluation from their senior peers to be stressful [28, 30]. Our manipulation of ego threat may not have been as powerful as others reported in the literature, as there was no physical presence from a known evaluator. Previous research, however, has consistently shown that the mere presence of a video camera is sufficient to cause evaluation stress [32–34]. The fact that the time pressure condition was perceived as stressful was not expected; however, this may reflect the specific wording used to introduce the condition. Subjects were asked to consider that, because of complications, some operating room procedures require quick completion. This instruction highlights the clinical relevance of the current training and may have provoked a more real life emotional response. Alternatively, asking trainees to better their best time during training provides a clearer understanding of the demands of the task and highlights the extent to which those demands might outweigh the trainees' perceived capabilities [5, 6].

The total workload data provide limited information beyond what could have been determined by asking subjects "how demanding was the task?" It provides no diagnostic information as to *why* multitasking and time pressure were the most demanding tasks (Fig. 2). This diagnostic information might be useful for a number of reasons. First, is the ability to assess why a procedure might be difficult, especially when performed under various demanding or stressful conditions (categorization). Second, the SURG-TLX may assist surgeons in making better decisions about the likely demands associated with introducing new techniques or technologies (e.g., robotic surgery) [13, 18, 20] into the operating room. In the ergonomics literature, where subjective workload is frequently considered during interface design, there has been a great deal of interest in understanding the "hidden" demands associated with the proliferation of technology [35, 36]. Third, the matching of appropriate training interventions to operator needs can only be assisted by diagnostic information about the sources of overload or stress. It is naïve to assume that the myriad of acute stress sources experienced by surgeons in the operating room will have an impact on performance through similar mechanisms. Training solutions should therefore be targeted at increasing coping resources for the particular demands experienced [37].

The current validation experiment followed the same approach as that adopted by a previous domain-specific adaptation of the TLX [23], by experimentally manipulating the demands of the task. Future research is required to assess "natural" sources of workload in the operating room for a variety of procedures and across experience levels. When possible, operators should complete both the paired comparison and the Likert scale components of the SURG-TLX. However, the Likert scale on its own can provide an informative visual analog of procedure demands. In this less stringent format the SURG-TLX has greater clinical utility; for example, it could be swiftly administered to help guide the self-reflection process of surgeons who have just performed poorly. Should the relative weighting between two dimensions remain unclear, paired comparisons could then be used to distinguish which of the dimensions makes the greatest contribution to workload.

Future research is required to determine whether the SURG-TLX is sensitive to the various combinations of stressors that occur in the operating room, and to the reflections of more experienced surgeons. However, this preliminary study supports the validity of the SURG-TLX as a multidimensional measure of surgical workload, which is sensitive to some of the typical stressors experienced during training.

Appendix 1: The SURG-TLX

Weighted rating

There are six rating scales which are meant for evaluating your experience during the procedure.

Please evaluate the procedure by marking "X" on each of the six scales at the point which best fits your experience. The scale ranges from "low" on the left to "high" on the right. Please read the descriptions carefully.

Mental Demands
How mentally fatiguing was the procedure?

Very Low   Very High

Physical Demands
How physically fatiguing was the procedure?

Very Low   Very High

Temporal Demands
How hurried or rushed was the pace of the procedure?

Very Low   Very High

Task Complexity
How complex was the procedure?

Not Very Complex   Very Complex

Situational Stress
How anxious did you feel while performing the procedure?

Not Very Anxious   Very Anxious

Distractions
How distracting was the operating environment?

Not Very   Very

Pairwise comparisons

Following are a set of titles listed into boxes within a grid. From these boxes, you will choose which title you deem more applicable to your experience of workload in the procedure.

Circle the title that you deem fitting of your experience.

Please consider your choices carefully and make them consistent with how you used the rating scales.

We are not looking for a right or wrong answer. We are only interested in your opinion.

Task Complexity **Or** **Mental Demand**	**Distractions** **Or** **Situational Stress**	**Task Complexity** **Or** **Distractions**
Task Complexity **Or** **Temporal demand**	**Mental demand** **Or** **Situational Stress**	**Physical Demand** **Or** **Distractions**
Mental demand **Or** **Physical demand**	**Situational Stress** **Or** **Physical demand**	**Situational Stress** **Or** **Task Complexity**
Temporal demand **Or** **Mental demand**	**Distractions** **Or** **Mental demand**	**Temporal demand** **Or** **Distractions**
Physical demand **Or** **Temporal demand**	**Physical demand** **Or** **Task Complexity**	**Temporal demand** **Or** **Situational Stress**

References

1. Arora S, Sevdalis N, Nestel D et al (2010) The impact of stress on surgical performance: a systematic review of the literature. Surgery 147:318–330
2. Wetzel CM, Kneebone RL, Woloshynowych M et al (2006) The effects of stress on surgical performance. Am J Surg 191:5–10
3. Sexton J, Thomas EJ, Helmreich RL (2000) Error, stress, and teamwork in medicine and aviation: cross sectional surveys. BMJ 320:745–749
4. Moorthy K, Munz Y, Dosis A et al (2003) The effect of stress-inducing conditions on the performance of a laparoscopic task. Surg Endosc 17:1481–1484
5. Lazarus RS, Folkman S (1984) Stress, appraisal, and coping. Springer-Verlag, New York
6. Lazarus RS (2000) Evolution of a model of stress, coping, and discrete emotions. In: Rice VH (ed) Handbook of stress, coping, and health: implications for nursing research, theory, and practice. Sage Publishing Inc, Thousand Oaks, pp 195–222
7. Marteau TM, Bekker H (1992) The development of a six-item short-form of the state scale of the Spielberger State Trait Anxiety Inventory (STAI). Br J Clin Psychol 31:301–306
8. Hancock PA, Desmond PA (2001) Stress, workload and fatigue, mahwah. Lawrence Erlbaum Associates, New Jersey
9. Eggemeier FT, Wilson GF, Kramer AF et al (1991) Workload assessment in multi-task environments. In: Damos DL (ed) Multiple-task performance. Taylor and Francis, Washington, DC, pp 207–216
10. Wierwille WA, Eggemeier FT (1993) Recommendations for mental workload measurement in a test and evaluation environment. Hum Factors 35:263–281
11. Prewett MS, Johnson RC, Saboe KN et al (2010) Managing workload in human–robot interaction: a review of empirical studies. Comput Hum Behav 26:840–856
12. DiDomenico A, Nussbaum MA (2008) Interactive effects of physical and mental workload on subjective workload assessment. Int J Ind Ergonom 38:977–983
13. Carswell CM, Clarke D, Seales WB (2005) Assessing mental workload during laparoscopic surgery. Surg Innov 12:80–90
14. Hart SG, Staveland LE (1988) Development of NASA-TLX (Task Load. Index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) Human mental workload. Amsterdam, Elsevier, pp 139–183
15. Muckler FA, Seven SA (1992) Selecting performance measures: objective versus subjective measurement. Hum Factors 34:441–455
16. Prabhu A, Smith W, Yurko Y et al (2010) Increased stress levels may explain the incomplete transfer of simulator-acquired skill to the operating room. Surgery 147:640–645
17. O'Connor A, Schwaitzberg SD, Cao CGL (2008) How much feedback is necessary for learning to suture? Surg Endosc 22:1614–1619
18. Sethi AS, Peine WJ, Mohammadi Y et al (2009) Validation of a novel virtual reality robotic simulator. J Endourol 23:503–508
19. Stefanidis D, Korndorffer JR Jr, Markley S et al (2007) Closing the gap in operative performance between novices and experts: does harder mean better for laparoscopic simulator training? J Am Coll Surg 205:307–313
20. Stefanidis D, Wang F, Korndorffer JR Jr et al (2010) Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. Surg Endosc 24:377–382
21. Karavidas MK, Lehrer PM, Lu S-E et al (2010) The effects of workload on respiratory variables in simulated flight: a preliminary study. Biol Psychol 84:157–160
22. Matthews R, Legg S, Charlton S (2003) The effect of cell phone type on drivers' subjective workload during concurrent driving and conversing. Accid Anal Prev 35:451–457
23. Pauzié A (2008) A method to assess the driver mental workload: the driving activity load index (DALI). IET Intell Transp Syst 2:315–322
24. Society of American Gastrointestinal and Endoscopic Surgeons (2010) Fundamentals of laparoscopic surgery (FLS). Available at http://www.flsprogram.org. Accessed August 25, 2010
25. Ritter EM, Scott DS (2007) Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery. Surg Innov 14:107–112
26. Scott DJ, Ritter EM, Tesfay ST et al (2008) Certification pass rate of 100% for fundamentals of laparoscopic surgery skills after proficiency-based training. Surg Endosc 22:1887–1893
27. Masters RSW, Lo CY, Maxwell JP et al (2008) Implicit motor learning in surgery: implications for multi-tasking. Surgery 143:140–145
28. Schuetz M, Gockel I, Beardi J et al (2008) Three different types of surgeon-specific stress reactions identified by laparoscopic simulation in a virtual scenario. Surg Endosc 22:1263–1267
29. Zheng B, Cassera MA, Martinec DV et al (2010) Measuring mental workload during the performance of advance laparoscopic tasks. Surg Endosc 24:45–50
30. Andreatta PB, Hillard M, Krain LP (2010) The impact of stress factors in simulation-based laparoscopic training. Surgery 147:631–639
31. Arora S, Sevdalis N, Aggarwal R et al (2010) Stress impairs psychomotor performance in novice laparoscopic surgeons. Surg Endosc 24:2588–2593
32. Gray R (2004) Attending to the execution of a complex sensorimotor skill: expertise differences, choking, and slumps. J Exp Psychol Appl 10:42–54
33. Lewis BP, Linder DE (1997) Thinking about choking? Attentional processes and paradoxical performance. Pers Soc Psychol B 23:937–944
34. Maxwell JP, Masters RSW, Poolton JM (2006) Performance breakdown in sport: the roles of reinvestment and verbal knowledge. Res Q Exerc Sport 77:271–276
35. Wickens CD (1992) Engineering psychology and human performance, 2nd edn. Harper-Collins, New York
36. Noyes JM, Garland KJ (2008) Computer- versus paper-based tasks: are they equivalent? Ergonomics 51:1352–1375
37. Arora S, Sevdalis N, Nestel D et al (2009) Managing intraoperative stress: what do surgeons want from a crisis training program? Am J Surg 197:537–543