Identifying Kinematic Markers Associated with Intraoperative Stress during Surgical Training Tasks

Yi Zheng[1] Grey Leonard[2] Juan Tellez[2] Herbert Zeh[2] Ann Majewicz Fey[1,2]

***Abstract*— Increased levels of stress can impair surgeon performance and patient safety during surgery. The aim of this study is to investigate the effect of short term stressors on laparoscopic performance through analysis of kinematic data. Thirty subjects were randomly assigned into two groups in this IRB-approved study. The control group was required to finish an extended-duration peg transfer task (6 minutes) using the FLS trainer while listening to normal simulated vital signs and while being observed by a silent moderator. The stressed group finished the same task but listened to a period of progressively deteriorating simulated patient vitals, as well as critical verbal feedback from the moderator, which culminated in 30 seconds of cardiac arrest and expiration of the simulated patient. For all subjects, video and position data using electromagnetic trackers mounted on the handles of the laparoscopic instruments were recorded. A statistical analysis comparing time-series velocity, acceleration, and jerk data, as well as path length and economy of volume was conducted. Clinical stressors lead to significantly higher velocity, acceleration, jerk, and path length as well as lower economy of volume. An objective evaluation score using a modified OSATS technique was also significantly worse for the stressed group than the control group. This study shows the potential feasibility and advantages of using the time-series kinematic data to identify the stressful conditions during laparoscopic surgery in near-real-time. This data could be useful in the design of future robot-assisted algorithms to reduce the unwanted effects of stress on surgical performance.**

I. INTRODUCTION

Performing surgery is stressful. Surgeons have to maintain continuous attention to detail while performing intricate tasks. Intraoperative stressors (Fig. 1) may include fatigue, disruptions, team work issues, time pressure, surgical complexity, high risk patients, and unexpected complications [1]. In addition, different types of surgery can be inherently more stressful to perform than others. For example, laparoscopic surgery has limitations in visualization, workspace volume, and an increased need for hand-eye coordination [2]–[4]. When it comes to robotic surgery, results are mixed in terms of measured surgeon stress levels using galvanic skin response when compared to either open surgery or virtual reality simulators, however, in neither study are the differences statistically significant [5], [6]. For complex motor tasks, it has been shown that external stressors can adversely affect motor performance [7]. The negative effects of stress on surgical performance include higher number of errors, less motion economy, and increased completion time [8]–[11].

*This work was supported in part by NSF grant number #1846726

[1]Yi Zheng and Ann Majewicz Fey are with the Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712 USA `yi.zheng@austin.utexas.edu`

[2]Grey Leonard, Juan Tellez, Herbert Zeh and Ann Majewicz Fey are with the Department of Surgery, the University of Texas Southwestern Medical Center, Dallas, TX 75390 USA `grey.leonard@utsouthwestern.edu`

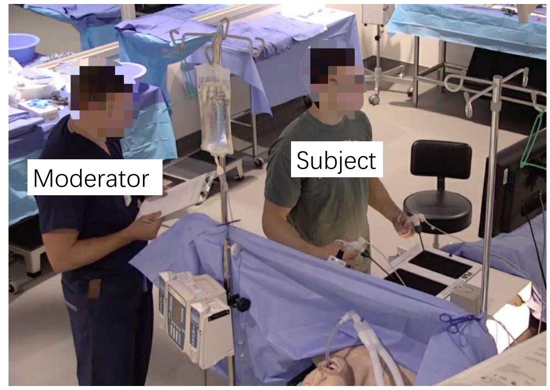


Fig. 1: Stresses in the operating room include both those associated with the patient status, as well as those associated with being a surgical trainee, who is directed and evaluated by an expert surgeon.

It has been shown that senior surgeons are able to develop stress management strategies that decrease the negative effect of stress on their performance, over time [12]–[15]. However; it is not fully understood how, specifically, those strategies change motor performance and how that information might be useful in the design of training platforms or feedback algorithms to detect and assist surgical trainees who experience stress while learning surgical tasks.

Physiological sensing is the most direct and traditional measure of stress (e.g., heart rate, skin conductance level). However, it also requires surgeons to wear sensors which could potentially interfere with surgeon's performance. In this study, we characterize the effect of clinical stress on surgical performance using a variety of kinematic metrics. Our long-term goal is to find kinematic markers associated with intraoperative stress that could be used to detect surgeon stress levels in real-time so as to mitigate the potential risk to the patient through the development of advanced control techniques on robotic-surgical platforms. This paper represents the first step towards that aim by better understanding the effects of stress on surgical movements.

II. BACKGROUND

Several studies have described how stress can effect surgical performance and there are a variety of sensors and analysis methods to measure physiological stress.

A. *Effect of Stress on Surgical Performance and Outcomes*

1) Performance Measurements: The Objective Structured Assessment of Technical Skill (OSATS) was developed and evaluated as a method of surgical technical skill assessment. The OSATS has shown its promise as a reliable method for testing operative skills in surgical trainees [16]. However, OSATS needs reviewer's rating and the resulting scores may be varied in different reviewers.

Alternatively, task-independent metrics (e.g., time, path length, smoothness, depth perception) extracted from the analysis of the laparoscopic instruments motions has been introduced as a technically sound approach for surgical performance assessment [17]. Several approaches of motion-tracking in laparoscopic surgery have been introduced, including electromagnetic sensors, optical and camera trackers. These studies demonstrated the potential feasibility of the kinematic data to access laparoscopic psychomotor skills using motion-based metrics such as path length, speed, or economy of volume [18]–[21]. The kinematic data was also adopted for evaluating surgical performance during robotics surgical training tasks, and it demonstrated the ability of objectively distinguishing between novice and expert performance as well as the training effects in the performance of training tasks [22].

2) The Effect of Stress on the Performance: Excessive levels of stress can compromise surgical performance [8]. The stressors led to impaired dexterity by showing an increased path length and a higher number of errors when the subject was under stressful conditions [9]. The cognitive distraction has been shown to have negative effects on the performance, such as a significantly greater time to task completion when subjects were distracted, and the overall score and economy of motion were negatively affected by distraction but didn't reach the level of statistical significance [10]. Furthermore, higher levels of stress correlated with increased completion time, lower economy of motion, and an increased number of errors [11]. However, none of these prior studies investigate pure kinematic metrics in depth.

B. *Tools and Techniques for Measuring Stress*

Traditional measures of stress includes self-report of stress level [2], [8], [12] and physiological sensing such as heart rate (HR) or heart rate variance (HRV), skin conductance level (SCL), and electrooculogram (EOG). Studies showed that all these physiological measures were increased by stressful conditions [2], [23]–[27]

However, self-report questionnaires are subjective and the physiological sensing systems are invasive since these technologies require wearable sensors which might interfere with subject's performance.

In this study, we would like to exploit the less invasive measurement - kinematic data - to identify the effect of stress during surgical training. In addition to being less invasive, kinematic data is inherently accessible on robotic laparoscopic surgery platforms, though maybe not yet easily available to research teams. Regardless, it has been shown that the kinematic data be used to predict expertise levels during training tasks on robotic-assisted surgical platforms [28]–[30]. By integrated the detection of stress with these robotic control platforms, there may be exciting opportunities to mitigate adverse effects of stress through robotic controls. This paper lays the groundwork for identifying kinematic markers of intraoperative stress.

III. Experimental Design

A. *Simulator Hardware*

1) FLS Trainer: The FLS (Fundamentals of Laparoscopic Surgery) trainer is a portable box trainer with a soft cover that can simulate the human abdomen. The trainer has 2 port holes for the laparoscopic instruments (Fig.2a) and a camera under the cover to simulate the laparoscopic camera and provides a field of vision (Fig.2c).

2) Electromagnetic Trackers: The electromagnetic trackers (Ascension 3D Guidance trakSTAR) were used to capture real-time data. The electromagnetic trackers were mounted to the handles of the tools using a pair of 3D printed adapters (Fig.2a) and used to obtain the x, y, z positions of the tool tips using a rigid body transformation and the geometry of the tools.

B. *Clinical Stressor*

In this study, the stressors included the vital signs from the monitor as well as the moderator's feedback. The vital signs are shown in Fig.3.

The moderator provided feedback to an illusory anesthesiologist and nurse circulator of the increased danger of the dummy patient and the need for adjunctive treatments such as intravenous fluids and blood transfusions (Fig.1) to simulate a busy and stressful operating room. Some feedback was directed at the participating subject to complete the task more quickly.

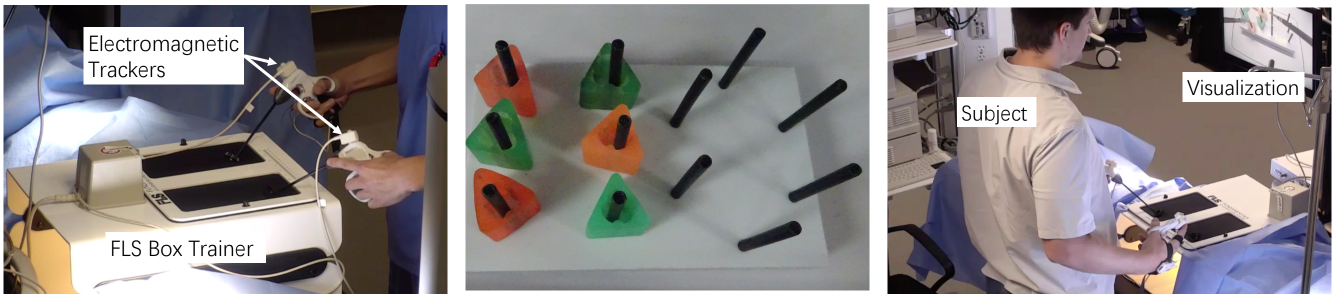
C. *Surgical Training Task*

The extended duration (e.g., 6 minutes) bimanual peg transfer drill was conducted using the FLS trainer as shown in Fig. 2b. The subjects were required to pick up the pegs and transfer them to another hand from one side on the board to another. The goal was to transfer as many blocks as possible whilst committing the fewest possible errors. Errors were defined as dropping a block or breaking a rule of transfer.

IV. Methods

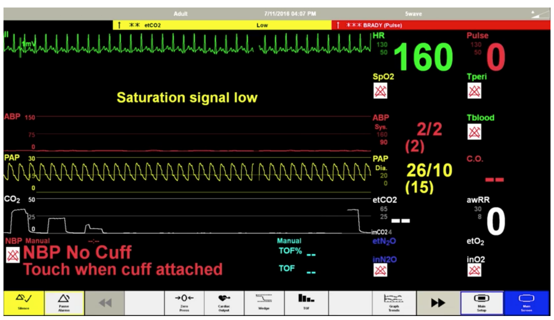
A. *Subject Recruitment*

Thirty users were recruited for this study. The subjects were medical students at the University of Texas Southwestern Medical Center in classes 1 through 4. Twenty-nine out of thirty participants were right-handed and 1 was left-handed. The study protocol was approved by UTD IRB office (UTD # 14-57). Participants had no previously reported muscular-skeletal injuries or diseases, or neurological disorders.

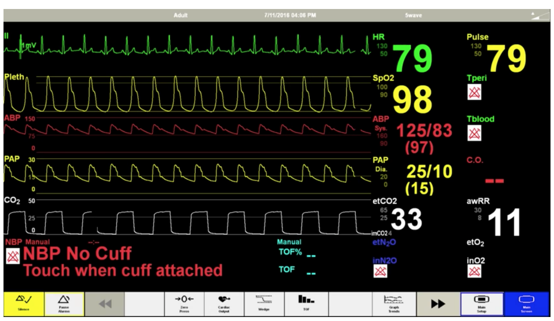


(a) Box Trainer and EM Trackers (b) FLS Peg Transfer Task (c) Subject and Visualization

Fig. 2: Simulator Setup.



(a) Normal Vital Signs



(b) Deteriorating Vital Signs

Fig. 3: Example normal and deteriorating vital signs during this experiment. Deteriorating vitals included increasing Heart Rate (HR), and decreasing Pulse, Blood Oxygen Level (SpO2), Blood Pressure (ABP), End-tidal CO2 (etCO2) and Respiratory Rate (awRR).

B. *Experimentation*

Each subject participated in several baseline surveys including a background questionnaire. Next, the subjects participated in a 10-minute tutorial on the fundamentals of laparoscopic (FLS) peg transfer drill to familiarize them with the instruments and with the requirements of the experimental task. In order to prevent bias, the subjects were randomized after the tutorial to the stressed or control group.

The experiment took place in a high-fidelity simulated operating room. The FLS peg transfer platform was placed in the abdominal section of a medical dummy which was draped. The vital sign monitor was in plain view. Several cameras recorded video from the experiment to capture images of the instrument tips and blocks, subject posture and the general environment.

The control group conducted the extended duration peg transfer task while hearing normal vital signs (Fig.3a) from the monitor for the duration of their task. The moderator did not provide any feedback on their performance. The stressed group performed under a period of progressively deteriorating vital signs (Fig.3b) with a particular increase in intensity beginning at the three-minute mark. The moderator also provided feedback to the stressed group and the feedback culminated in 30 seconds of cardiac arrest and the expiration of the dummy patient, occurring simultaneously with the end of stressed six-minute task.

C. *Data Analysis*

All objective metrics of performance were based on the kinematic features of the tool tips. The tip positions were calculated using EM tracker positions and a rigid body transformation using the tool geometry.

1) Data Acquisition: The kinematic data was streamed and recorded from the EM trackers through ROS topics [31]. In this study, the kinematic features including the x-, y-, z- positional coordinates in space and quaternions x-, y-, z-, w- were collected. The positional coordinates determined the tool positions in space and the quaternions were used to determine the rotation matrix for calculating the 3 dimensional tool tip positions ($P = [Px, Py, Pz]^T$).

2) Data Processing: The kinematic data was recorded at a frequency of 256 Hz from the EM trackers. In order to reduce the noise and improve computational efficiency, after calculating the tool tip positions, we used an established method to down-sample the kinematic data to 5 Hz using a cubic spline to enforce a constant sampling rate between data points, therefore, smoothed the data for kinematic metrics calculation [22].

3) Kinematic Metrics: The kinematic metrics included velocity (V), acceleration (A), jerk (J), Path Length (PL) and the Economy of Volume (EV) of the tool tip.

The Velocity (V) was time series data calculated as follow:

$$V_t = \frac{\sqrt{(P_{t+1} - P_t)^T (P_{t+1} - P_t)}}{T_{t+1} - T_t} \tag{1}$$

P_t is the 3-D position at time t, and T_t is the time stamp at time t.

The Acceleration (A) and the jerk (J) were time series data calculated in the similar way:

$$A_t = \frac{V_{t+1} - V_t}{T_{t+1} - T_t}, J_t = \frac{A_{t+1} - A_t}{T_{t+1} - T_t} \quad (2)$$

The Path Length (PL) is the sum of the displacement at each time point and it indicates the total length traveled. This parameter describes the spatial distribution of the tip of the laparoscopic instrument in the work space of the task. A compact "distribution" is characteristic of an expert [17]:

$$PL = \sum_{T_{start}}^{T_{end}} \sqrt{(P_{t+1} - P_t)^T (P_{t+1} - P_t)} \quad (3)$$

The Economy of Volume (EV) is a single-value data indicating the efficiency of occupying the space [20], and a larger value of EV indicates a better performance:

$$EV = \frac{\sqrt[3]{(x_{max} - x_{min})(y_{max} - y_{min})(z_{max} - z_{min})}}{PL} \quad (4)$$

4) Video Review: Besides the kinematic metrics, video review was conducted to include measurement of the counts of blocks transferred (N) and errors committed (Er). Additionally, a blinded, independent reviewer with training in OSATS [16] scoring graded each subject using a modified OSATS (mOSATS) rubric. The subsections included in scoring were respect for tissue (RFT), time and motion (TM), instrument handling (IH) and the total score (TOT). Each of these scores ranged from 1 to 5, with 5 representing the best and 1 the worst performance.

D. Analysis Methods

We examined the distribution properties of all the metrics mentioned above.

The time series data of Velocity, Acceleration and Jerk were non-gaussian distributed while the other metrics (statics data) such as Path Length, Economy of Volume and mOSATS scores were gaussian distributed.

According to Section IV-B, the experiment length was 6 minutes and the stressed group was experiencing the clinical stress which progressively increased its intensity at 3-minute mark and culminated at the end of the task. Therefore, we divided the collected data into two halves (H1 vs. H2) and ideally, the stress should show more effect in the second half. Therefore, in order to study the significant effect of the stress, we first compared the data of the second half between Control and Stressed Groups, then the data of Stressed Group between First and Second Halves.

Therefore, according to data distribution properties and data dependencies, as summarized in Table I, we used different methods for statistical analysis.

Comparisons	Statistical Analysis Methods	Applied to
Time Series Data between Groups (Non-Gaussian distributed, independent)	Mann-Whitney U-test	Table II
Static Data between Groups (Gaussian distributed, independent)	Independent t-test	Table IV
Time Series Data between Halves in Stressed Group (Non-Gaussian distributed, dependent)	Wilcoxon signed rank test	Table III
Static Data between Halves in Stressed Group (Gaussian distributed, dependent)	Dependent t-test	Table V

TABLE I: Summary of the statistical analysis methods for different data.

V. Results

A. Control Group vs. Stressed Group

The results of comparisons between stressed and control groups are shown in Table II, Table IV, Fig 4 and Fig 5.

For kinematic metrics, stressed group has greater Velocity (Non-dominant Hand: $p < 0.0125$, Dominant Hand: $p < 0.0001$), Acceleration (Non-dominant Hand: $p = 0.0396$, Dominant Hand: $p = 0.0016$), and Jerk (Non-dominant Hand: $p < 0.0001$, Dominant Hand: $p < 0.0001$) than control group for both hands. However, Path Length (Non-dominant Hand: $p = 0.9772$, Dominant Hand: $p = 0.6467$) and Economy of Volume (Non-dominant Hand: $p = 0.2434$ vs. Dominant Hand: $p = 0.6596$) cannot show significant difference between groups.

For mOSATS scores, control group has greater score values than stressed group in metrics Respect for Tissue (RFT: $p = 0.0198$), Instrument Handling (IH: $p = 0.0158$) and the Total Score (TOT: $p = 0.0067$) which means better performance in control group.

Even though metrics of path length, economy of volume, number of blocks, number of errors and mOSATS-TM score cannot show significance between groups, the desired trend can be found, i.e. more blocks transferred (Fig.5c), less errors made (Fig.5d), and greater mOSATs scores (Fig.5e) in the control group.

B. First Half vs. Second Half of Stressed Group

We also studied the effect of the intensity of stress. We analyzed the performance of stressed group between the first and the second half of the experiment, as shown in Table III, Table V, Fig 4 and Fig 5.

The effect of increasingly intensive stress are significant. The second half which is with more intensive stress, has greater Velocity, Acceleration, Jerk, Path Length and lower Economy of Volume.

mOSATS scores between the two halves also support the kinematic metrics. The first half has significantly greater score values than the second half in metrics Respect for Tissue (RFT), Instrument Handling (IH) and the Total Score (TOT).

However, the mOSATS-TM (Time and Motion) metrics failed to show significant results in evaluating the subject movement as shown in Table IV and Table V. Therefore,

Metrics	Hand	Control vs. Stressed (median[IQR])	p
V	ND	0.0249[0.0329] <0.0260[0.0329]	0.0125
	D	0.0218[0.0327] <0.0229[0.0340]	<0.0001
A	ND	0.1495[0.1854] <0.1526[0.1918]	0.0396
	D	0.1222[0.1793] <0.1245[0.1832]	0.0016
J	ND	1.1856[1.3909] <1.2326[1.4681]	<0.0001
	D	0.9411[1.3196] <0.9666[1.3780]	<0.0001

TABLE II: Comparison of Velocity, Acceleration and Jerk between Control and Stressed Groups using Mann Whitney U-test. ND: Nondominant Hand, D: Dominant Hand.

Metrics	Hand	Stressed H1 vs. H2 (median[IQR])	p
V	ND	0.0236[0.0302] <0.0287[0.0358]	<0.0001
	D	0.0209[0.0308] <0.0253[0.0373]	<0.0001
A	ND	0.1361[0.1754] <0.1715[0.2072]	<0.0001
	D	0.1107[0.1583] <0.1446[0.2065]	<0.0001
J	ND	1.1037[1.3493] <1.3758[1.5879]	<0.0001
	D	0.8543[1.1984] <1.1193[1.5584]	<0.0001

TABLE III: Comparison of Velocity, Acceleration and Jerk between First and Second Halves in Stressed groups using Wilcoxon signed rank test. ND: Nondominant Hand, D: Dominant Hand.

according to our analysis, the kinematic metrics show potential advantages for evaluating the effect of stress over the method of mOSATS.

VI. Discussion

Prior work has shown that experts have significantly greater velocity than pre-trained novices which means a better performance [22]. However, according to our results, mOSATS successfully showed that the stressed group had worse performance as well as greater velocities. Some limitations of our study is that this was a simple peg transfer task performed by medical students, meaning our results lack a wide range of expertise levels. Regardless, our results suggest the importance of further investigate the role of velocity in detecting both stress and expertise. Future studies with more complicated surgical training tasks and subjects of different expertise levels should be conducted to better interpret the underlying properties of movement velocity.

Metrics	Hand	Control vs. Stressed (mean(SD))	p
PL	ND	Not Significant	0.9772
	D	Not Significant	0.6467
EV	ND	Not Significant	0.2434
	D	Not Significant	0.6596
mOSATS -RFT		3(1.1767) >2.2(0.4140)	0.0198
mOSATS -TM		Not Significant	0.1250
mOSATS -IH		1.8571(1.0271) >1.1333(0.3519)	0.0158
mOSATS -TOT		7.6429(2.5603) >5.5333(1.0601)	0.0067
# of Blocks		Not Significant	0.9234
# of Errors		Not Significant	0.6522

TABLE IV: Comparison of Static Metrics (Path Length, Economy of Volume, mOSATS scores, number of blocks transferred, number of errors made) between Control and Stressed groups using independent t-test. ND: Nondominant Hand, D: Dominant Hand.

Metrics	Hand	Stressed H1 vs. Stressed H2 (mean(SD))	p
PL	ND	5.5895(1.4606) <6.6961(1.3887)	<0.0001
	D	5.3662(1.1530) <6.4489(1.1453)	<0.0001
EV	ND	0.0166(0.0033) >0.0142(0.0019)	0.0014
	D	0.0200(0.0048) >0.0152(0.0040)	0.0035
mOSATS -RFT		3.2000(1.0823) >2.2000(0.4140)	0.0017
mOSATS -TM		Not Significant	0.2620
mOSATS -IH		1.6667(0.7237) >1.1333(0.3519)	0.0148
mOSATS -TOT		7.3333(1.6762) >5.5333(1.0601)	<0.0001

TABLE V: Comparison of Static Metrics (Path Length, Economy of Volume, mOSATS scores) between first and second halves in stressed group using dependent t-test. ND: Nondominant Hand, D: Dominant Hand.

Our results also agree with prior work that lower jerk values describe a better performance [17]. The metric of economy of volume failed to show significant difference between control and stressed groups which is consistent with prior results that motion economy didn't have statistical significance between distracted and undistracted groups [10].

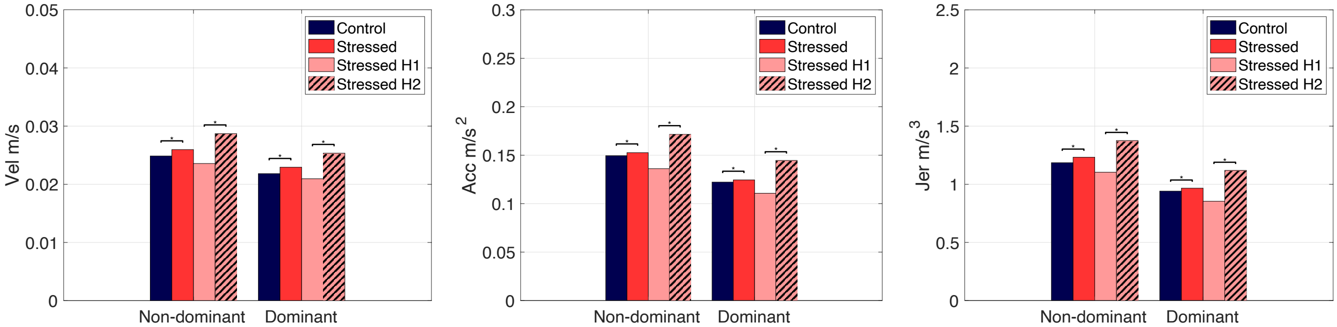
VII. Conclusion

In this study, we exposed subjects to commonly experienced clinical stressors during surgical operation. Our results show that both kinematic metrics and mOSATS scores showed significant differences between the control and stressed groups. The clinical stressors had a negative effect on surgical performance, as measured by the mOSATS scores, and our kinematic metrics of velocity, acceleration, jerk, path length, and ecomony of volume are also negatively impacted by stress conditions for both the dominant and nondominant hands. To be more specific, the stressed group's movement is less smooth but faster than the control group. Overall the stress group resulted in lower mOSATS scores and the control group had better performance in treating the tissue, and handling and moving with the instruments relative to the stressed group.

We also found the shortcomings of using mOSATS to evaluate the effect of surgical stress. The metric of mOSATS-TM, which was designed to assess the subject motion during surgical training, wasn't able to evaluate the effect of stress on subject movement. Kinematic metrics show an advantage in evaluating the effect on movement over the mOSATS dimensions alone.

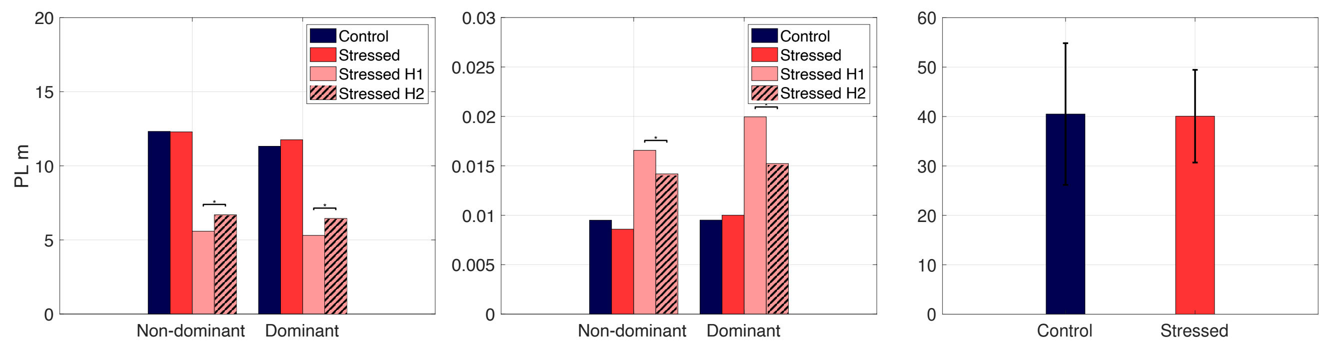
Since both methods of mOSATS and kinematic analysis can evaluate the performance under stress conditions, future studies should investigate the correlation between mOSATS scores and different kinematic metrics. And potentially, finding novel metrics which can respectively interpret mOSATS scores could enable automatic collection for mOSATS scores from kinematic features.

This study can serve as the groundwork for future work to providing preventative control strategies to reduce the unwanted effect of stress during surgical training, consequently, improving surgical training outcomes and patient safety. In future work, we will implement the real-time detection of

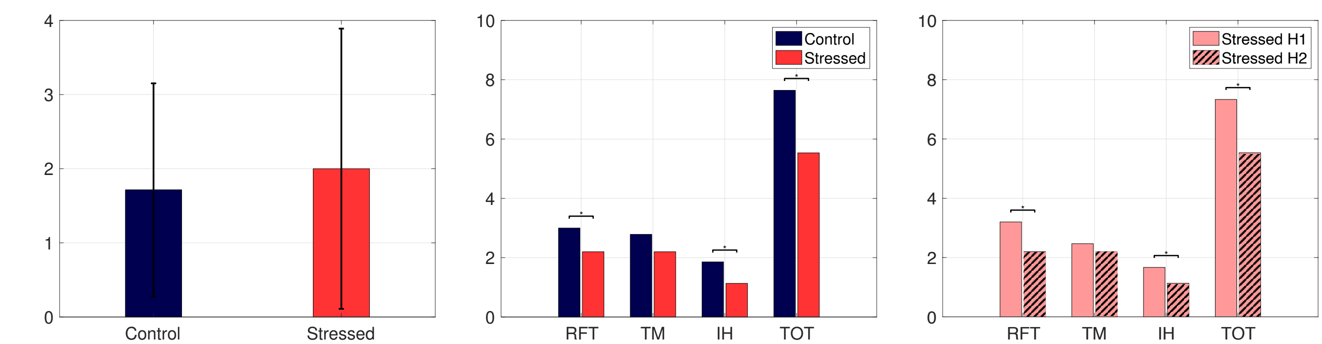


(a) Velocity (b) Acceleration (c) Jerk

Fig. 4: Kinematic Results: (a) Velocity, (b) Acceleration, (c) Jerk, the group medians and significant differences were plotted. The horizontal bar and star indicate the significance. ND: Non-dominant Hand, D: Dominant Hand. H1 and H2 correspond to the first half (H1, 3min) and second half (H2, 3 min) for the stressed group. Brackets indicate significance.



(a) Path Length (b) Economy of Volume (c) Blocks Transferred



(d) Errors Made (e) mOSATS Scores between Groups (f) mOSATS Scores between H1 and H2

Fig. 5: Static Results Summary: Group means were illustrated. (a) Path Length, (b) Economy of Volume (c) Blocks Transferred, (d) Errors Made, (e) mOSATS dimensions between groups and (f) mOSATS dimensions before the first (H1) and second (H2) half of the experiment. The brackets indicate significance.

experienced stress using kinematic data. The detection of stress could trigger haptic cues on robotic-assisted surgery platform to provide stress coping strategies, such as pausing and slowing down, to mitigate the negative effect of excessive stress [1].

ACKNOWLEDGMENT

The authors would like to thank the UTSW Training Resident Doctors As Innovators in Science (TARDIS) program, supporting co-Author Grey Leonard. We also thank the UTSW Simulation Center director and staff for their invaluable assistance with data collection.

References

[1] N. E. Anton, P. N. Montero, L. D. Howley, C. Brown, and D. Stefanidis, “What stress coping strategies are surgeons relying upon during surgery?” in *American Journal of Surgery*, vol. 210, no. 5. Elsevier Inc., nov 2015, pp. 846–851.

[2] R. Berguer, W. D. Smith, and Y. H. Chung, “Performing laparoscopic surgery is significantly more stressful for the surgeon than open surgery,” *Surgical Endoscopy*, vol. 15, no. 10, pp. 1204–1207, 2001.

[3] R. Berguer, G. T. Rab, H. Abu-Ghaida, A. Alarcon, and J. Chung, “A comparison of surgeons’ posture during laparoscopic and open surgical procedures,” *Surgical Endoscopy*, vol. 11, no. 2, pp. 139–142, 1997.

[4] A. G. Gallagher, N. McClure, J. McGuigan, K. Ritchie, and N. P. Sheehy, “An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills,” *Endoscopy*, vol. 30, no. 7, pp. 617–620, 1998.

[5] R. Berguer and W. Smith, “An Ergonomic Comparison of Robotic and Laparoscopic Technique: The Influence of Surgeon Experience and Task Complexity,” *Journal of Surgical Research*, vol. 134, no. 1, pp. 87–92, jul 2006.

[6] Z. Wang, M. Kasman, M. Martinez, R. Rege, H. Zeh, D. Scott, and A. M. Fey, “A comparative human-centric analysis of virtual reality and dry lab training tasks on the da vinci surgical platform,” *Journal of Medical Robotics Research*, vol. 4, no. 03n04, p. 1942007, 2019.

[7] E. D. Ryan, “Effects of stress on motor performance and learning,” *Research Quarterly. American Association for Health, Physical Education and Recreation*, vol. 33, no. 1, pp. 111–119, 1962.

[8] S. Arora, N. Sevdalis, D. Nestel, T. Tierney, M. Woloshynowych, and R. Kneebone, “Managing intraoperative stress: what do surgeons want from a crisis training program?” *American Journal of Surgery*, vol. 197, no. 4, pp. 537–543, apr 2009.

[9] K. Moorthy, Y. Munz, A. Dosis, S. Bann, and A. Darzi, “The effect of stress-inducing conditions on the performance of a laparoscopic task,” *Surgical Endoscopy and Other Interventional Techniques*, vol. 17, no. 9, pp. 1481–1484, sep 2003.

[10] K. H. Goodell, C. G. Cao, and S. D. Schwaitzberg, “Effects of cognitive distraction on performance of laparoscopic surgical tasks,” *Journal of Laparoendoscopic and Advanced Surgical Techniques*, vol. 16, no. 2, pp. 94–98, apr 2006.

[11] S. Arora, N. Sevdalis, R. Aggarwal, P. Sirimanna, A. Darzi, and R. Kneebone, “Stress impairs psychomotor performance in novice laparoscopic surgeons,” *Surgical Endoscopy*, vol. 24, no. 10, pp. 2588–2593, 2010.

[12] C. M. Wetzel, R. L. Kneebone, M. Woloshynowych, D. Nestel, K. Moorthy, J. Kidd, and A. Darzi, “The effects of stress on surgical performance,” *American Journal of Surgery*, vol. 191, no. 1, pp. 5–10, jan 2006.

[13] K. Moorthy, Y. Munz, D. Forrest, V. Pandey, S. Undre, C. Vincent, and A. Darzi, “Surgical crisis management skills training and assessment: A stimulation-based approach to enhancing operating room performance,” *Annals of Surgery*, vol. 244, no. 1, pp. 139–147, jul 2006.

[14] A. Yamamoto, T. Hara, K. Kikuchi, T. Hara, and T. Fujiwara, “Intraoperative stress experienced by surgeons and assistants,” *Ophthalmic Surgery and Lasers*, vol. 30, no. 1, pp. 27–30, jan 1999.

[15] C. M. Wetzel, S. A. Black, G. B. Hanna, T. Athanasiou, R. L. Kneebone, D. Nestel, J. H. Wolfe, and M. Woloshynowych, “The effects of stress and coping on surgical performance during simulations,” *Annals of surgery*, vol. 251, no. 1, pp. 171–176, 2010.

[16] J. A. Martin, G. Regehr, R. Reznich, H. Macrae, J. Murnaghan, C. Hutchison, and M. Brown, “Objective structured assessment of technical skill (OSATS) for surgical residents,” *British Journal of Surgery*, vol. 84, no. 2, pp. 273–278, feb 1997.

[19] S. Yamaguchi, D. Yoshida, H. Kenmotsu, T. Yasunaga, K. Konishi, S. Ieiri, H. Nakashima, K. Tanoue, and M. Hashizume, “Objective assessment of laparoscopic suturing skills using a motion-tracking system,” *Surgical Endoscopy*, vol. 25, no. 3, pp. 771–775, 2011.

[17] S. Cotin, N. Stylopoulos, M. Ottensmeyer, P. Neumann, D. Rattner, and S. Dawson, “Metrics for Laparoscopic Skills Trainers: The Weakest Link!” in *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002*, T. Dohi and R. Kikinis, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 35–43.

[18] V. Datta, S. Mackay, M. Mandalia, and A. Darzi, “The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model,” *Journal of the American College of Surgeons*, vol. 193, no. 5, pp. 479–485, nov 2001.

[20] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, Á. Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo, and E. J. Gómez, “EVA: Laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment,” *Surgical Endoscopy*, vol. 27, no. 3, pp. 1029–1039, oct 2013.

[21] J. A. Sánchez-Margallo, F. M. Sánchez-Margallo, ·. I. Oropesa, S. Enciso, and E. J. Gómez, “Objective assessment based on motion-related metrics and technical performance in laparoscopic suturing,” *Int J CARS*, vol. 12, pp. 307–314, 2017.

[22] T. N. Judkins, D. Oleynikov, and N. Stergiou, “Objective evaluation of expert and novice performance during robotic surgical training tasks,” *Surgical Endoscopy and Other Interventional Techniques*, vol. 23, no. 3, pp. 590–597, mar 2009.

[23] B. Böhm, N. Rötting, W. Schwenk, S. Grebe, and U. Mansmann, “A prospective randomized trial on heart rate variability of the surgical team during laparoscopic and conventional sigmoid resection,” *Archives of Surgery*, vol. 136, no. 3, pp. 305–310, 2001.

[24] E. Czyzewska, K. Kiczka, A. Czarnecki, and P. Pokinko, “The surgeon’s mental load during decision making at various stages of operations,” *European Journal of Applied Physiology and Occupational Physiology*, vol. 51, no. 3, pp. 441–446, 1983.

[25] A. P. Tendulkar, G. P. Victorino, T. J. Chong, M. K. Bullard, T. H. Liu, and A. H. Harken, “Quantification of surgical resident stress ”on call”,” *Journal of the American College of Surgeons*, vol. 201, no. 4, pp. 560–564, oct 2005.

[26] W. Boucsein, *Electrodermal activity: Second edition*. Springer US, aug 2012, vol. 9781461411.

[27] W. D. Smith, Y. H. Chung, and R. Berguer, “A virtual instrument ergonomics workstation for measuring the mental workload of performing video-endoscopic surgery,” in *Studies in Health Technology and Informatics*, vol. 70. IOS Press, 2000, pp. 309–315.

[28] Z. Wang and A. Majewicz Fey, “Deep learning with convolutional neural network for objective skill evaluation in robot-assisted surgery,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, no. 12, pp. 1959–1970, dec 2018.

[29] L. Tao, E. Elhamifar, S. Khudanpur, G. D. Hager, and R. Vidal, “Sparse hidden Markov models for surgical gesture classification and skill evaluation,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7330 LNCS. Springer, Berlin, Heidelberg, 2012, pp. 167–177.

[30] M. J. Fard, S. Ameri, R. Darin Ellis, R. B. Chinnam, A. K. Pandya, and M. D. Klein, “Automated robot-assisted surgical skill evaluation: Predictive analytics approach,” *International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 14, no. 1, feb 2018.

[31] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Ng, “Ros: an open-source robot operating system,” vol. 3, 01 2009.