

OBJECTIVE MEASUREMENT FOR SURGICAL SKILL EVALUATION

Siti Nor Zawani Ahmmad^a, Eileen Su Lee Ming^{a*}, Yeong Che Fai^b,
Suneet Sood^c, Anil Gandhi^c

^aFaculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

^bCentre of Artificial Intelligence and Robotics (CAIRO), Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

^cFaculty of Medicine, Jeffrey Cheah School of Medicine and Health Sciences, Monash University, 80100 Johor Bahru, Malaysia

Article history

Received

30 October 2015

Received in revised form

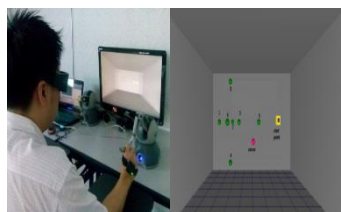
15 February 2016

Accepted

25 February 2016

*Corresponding author
eileensu@utm.my

Graphical abstract



Abstract

The purpose of this study was to identify measurable parameters that can be used to quantitatively assess psychomotor skills, specifically for surgical skills assessment. Sixteen participants were recruited from two groups: surgeon (N = 5) and non-surgeon (N = 11). Both groups underwent a psychomotor test using a custom developed 'Green Target' module which was designed using a virtual reality system. Six parameters were used to compare the psychomotor skills between the two groups. The results showed that surgeons outperformed the non-surgeons in five out of six parameters investigated and the difference was statistically significant. The average normalised comparison values for surgeons and non-surgeons for motion path accuracy, motion path precision, economy of movement, end-point accuracy and end-point precision were 0.13 ± 0.12 and 0.17 ± 0.12 , 0.08 ± 0.11 and 0.10 ± 0.10 , 3.76 ± 1.76 and 4.08 ± 2.24 , 0.12 ± 0.10 and 0.17 ± 0.11 , 0.04 ± 0.10 and 0.07 ± 0.10 respectively, $p < 0.05$). These parameters can potentially be used to objectively assess the performance of surgical skill.

Keywords: Psychomotor skills, assessment parameters, computer based measurements

Abstrak

Tujuan kajian ini adalah untuk mengenal pasti parameter yang boleh diukur dan boleh digunakan untuk menilai kemahiran psikomotor secara kuantitatif, khusus untuk penilaian kemahiran pembedahan. Enam belas peserta telah direkrut dari dua kumpulan: pakar bedah (N = 5) dan bukan pakar bedah (N = 11). Kedua-dua kumpulan telah menjalani ujian psikomotor menggunakan modul 'Sasaran Hijau' yang direka menggunakan sistem realiti maya. Enam parameter telah digunakan untuk membandingkan kemahiran psikomotor antara kedua-dua kumpulan. Hasil kajian menunjukkan bahawa pakar bedah mengatasi bukan pakar bedah dalam lima daripada enam parameter yang disiasat dan perbezaan secara statistik adalah ketara. Purata nilai perbandingan normal untuk pakar bedah dan bukan pakar bedah bagi akurasi pergerakan jalan, presisi pergerakan jalan, ekonomi pergerakan, akurasi titik akhir dan presisi titik akhir adalah 0.13 ± 0.12 dan 0.17 ± 0.12 , 0.08 ± 0.11 dan 0.10 ± 0.10 , 3.76 ± 1.76 dan 4.08 ± 2.24 , 0.12 ± 0.10 dan 0.17 ± 0.11 , 0.04 ± 0.10 dan 0.07 ± 0.10 masing-masing, $p < 0.05$). Parameter ini boleh berpotensi digunakan untuk menilai secara objektif prestasi kemahiran pembedahan.

Kata kunci: Kemahiran psikomotor, parameter penilaian, ukuran berasaskan komputer

© 2016 Penerbit UTM Press. All rights reserved

1.0 INTRODUCTION

The main goal of surgical skills assessment is to identify surgeons who can operate safely and skilfully. In addition to sufficient medical knowledge, technical skills and dexterity play important roles in determining the outcome of surgery. Conventionally, learning of surgical skills relies heavily on apprentice-style training 'See one, do one, teach one' [1] and normally uses live patients. However, this method has drawbacks due to concerns regarding patient safety, time constraint [2] and financial pressure on hospitals due to the increasing insurance cost of malpractice, surgical equipment and cost of training residents [3-4]. In the process of acquiring new skills in the operating room (OR), trainees may expose patients to harm because of their lack of experience and technical skills. Besides that, due to time constraints, trainees may have less opportunity to learn and practice under supervision. Therefore, training of surgical skills should ideally be done outside of the operating theatre in an unhurried and non-threatening environment, until a baseline surgical skills have been achieved before they perform an actual procedure [5].

Various assessment methods have been developed and validated to assess surgical skills. However, most of the assessments lack objective and quantitative method. Assessment using operative log book is very commonly used in United Kingdom. Generally, the operative log book records the experiences that trainee has gained through the training, and will be submitted during annual assessment or at the time of examination [6]. Even though it records the experience when performing of the surgical procedure, it does not reflect the technical skill. The lack of information about skill proficiency indicates that this method is subjective and has poor validity. Assessments based on clinical outcomes, and mortality and morbidity data are usually used to indicate the proficiency of surgical skills [7-8]. However, mortality and morbidity data are strongly influenced by many additional factors such as patient characteristic and the complication of a case [9]. Thus, it is not accurate to identify the surgeon's performance level based on mortality and morbidity. Besides, it takes a very long period and involves a massive number of patients for the data to produce significant results.

The expert surgeon typically observes the trainee performance in the operating room and provides verbal feedback. The performance measure may be greatly influenced by the observer's personal judgment. In the process of evaluation, the expert rates the trainee's performance by using specific assessment criteria such as checklist and global rating scale [10-11]. A checklist provides a list of steps to complete the specific task that the trainee should perform. When comparing the performance between surgeons and novices using the checklist alone, the examiners usually are unable to identify whether candidates have used appropriate steps or

not, even though they finished the task well [12]. Many published works have reported that the checklist and GRS show good reliability and validity [13-5]. However, these type of assessment are still potentially affected by inter and intra-rater variability and lead to recall bias. Furthermore, specific skills such as psychomotor skill cannot be quantified using observation.

With emerging new technology, computational based systems using motion analysis and virtual reality system have been proposed for surgical assessment [16-18]. These methods are more objective and provide quantitative measurements compared to structured human grading. By using simulation on computer based systems, training and assessment can happen together, where collected data from training can be used and analysed to provide quantitative assessment [19-20]. Most of quantitative assessment methods for surgical skill require a set of measurement parameters. There are several performance parameters that have been studied previously such as time to complete the task, economy of movement, accuracy, motion smoothness and force variability [21-23]. Measurement parameters which are able to differentiate between expert and non-expert significantly help the construction and validation of an assessment tool. More importantly, with the help of computerized-based assessment parameters, trainees obtain immediate feedbacks to improve their performance.

The goal of this study is to investigate a set of assessment parameters to assess psychomotor skill using a custom developed module based on reaching and pointing tasks. The assessment parameters were analysed from the extraction of motion data during experimental task. Then, the parameters were used to differentiate the performance between surgeon and non-surgeon groups.

2.0 EXPERIMENTAL

2.1 Subject

Sixteen subjects without known hand pathology participated in this study and they were divided into two groups: surgeon and non-surgeon. Surgeon's group comprised five surgeons (female=1, male=4) with at least 3 years' experience in surgery and aged between 33 and 42 years' old. Non-surgeon's group comprised eleven healthy adults (female=5, male=6) aged between 22 and 26 years' old. All subjects were right-hand dominant. The details of the nature and purpose of the research were explained to them before informed consent was obtained from the participants.

2.2 Experimental Set-up

The PHANTOM Omni haptic device from Sensable Technologies was used in this study for position measurement during movement. The haptic feedback loop ran at 1000 Hz. This haptic device provided 6 degree of freedom (DOF) positional and orientation sensing using digital encoders with nominal accuracy of 0.055 mm. Besides that, 3-DOF force feedback can be provided with continuous force of 0.88 N and maximum force of 3.3 N within 160x120x70mm³.

The basic framework of user interface for the task module was developed using Microsoft Visual Studio C++ while the graphics of virtual environment and objects was developed using OpenGL library. The user interface was able to display the motion of the haptic stylus. For this study, visual display was provided through a 3D monitor Acer HS244HQ with a pair of active 3D shutter glasses (built-in IR emitter). The 3D monitor also has 23.6 inch display with 1920x1080 pixel full HD resolution at 120Hz refresh rate. The graphics card used together with this study is the nVidia GeForce 54m series.

An experimental software module was developed to investigate subjects' movement. During the experiment, subjects were able to perceive their movements in 3-dimension, and could visually estimate the depth of their movements using a pair of shutter glasses and the 3D feature of the monitor screen. The origin of the OpenGL co-ordinate {x,y,z} was located at the centre of the computer screen. The positive of the x-axis pointed to the right, the y-axis pointed upwards while the z-axis pointed towards the viewer.

In this study, each subject performed sets of experiment using their right hand. Visual display contained a static purple sphere inside a yellow box, located at the right side of the screen. This acted as a starting point and a pink sphere, which can move freely corresponding to the movement of the phantom's stylus tip position, acted as the cursor. In the middle region, there were 7 green spheres, which represented the targets for the subject to aim at. They were located at different horizontal, vertical and depth planes. The 7 green targets were presented one at a time (Figure 1). Subjects needed to grip the stylus of the PHANTOM with their hands between fingers and thumb as if they were holding a pen. When the subject was ready to start the experiment, a keyboard press set the pink sphere to overlap the starting point. Simultaneously, the green target appeared and data collection began. The pink sphere was held at the starting point with haptic force for 1s until an indicator "Go" appeared on screen. Then, the convergent force at starting point was switched off and subjects were free to start their trajectory towards the target. No time limit was imposed on the subject. After reaching the target point, they needed to align and hold the cursor at the target as accurately as possible for 3 s. All subjects were asked to complete two sessions of the

experiment. For each session, each subject repeated the movement three times for every target. Hence, the total trials for each subject to complete the experiment were 42 trials (7 targets x 3 repetitions x 2 sessions).

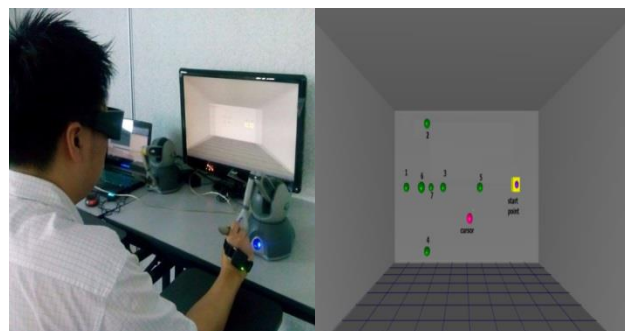


Figure 1 Experimental setup (left) and 2D screenshot of visual feedback during the experiment (right). Subjects were required to move the cursor (pink ball) from start point (yellow box) to green targets (green balls)

2.3 Experimental Data Analysis

Based on collected data, several useful parameters were extracted and processed in Matlab software (The Mathworks, USA). The captured data were separated into two sets: reaching data when subjects were moving towards the target and pointing data when subjects reached their target. The reaching data was used to analyse dynamic movements while the pointing data was used to analyse static accuracy of subjects. For reaching analysis, extraction of data started from the moment subjects moved their cursor until they reached the target point. For pointing analysis, data was extracted within three-second time frame when subjects reached the target and kept the cursor on the target for 3s. The parameters analysed in reaching analysis were motion path accuracy and precision, economy movement and motion smoothness. For pointing analysis, two parameters were analysed: end-point accuracy and end-point precision.

Motion path accuracy was identified by calculating the average of all the deviation errors throughout their trajectories. Deviation error, d represented the error made by the subjects compared to the ideal trajectory (Figure 2). The magnitude of the deviation error, d was computed by calculating the shortest distance between the ideal trajectory with the cursor point. Smaller values of mean deviation error indicated higher accuracy because subjects were able to make a trajectory close to ideal.

Motion path precision was identified by calculating the standard deviation of all the deviation errors, d throughout the trial. A smaller value of standard deviation indicated higher precision because

subjects were able to maintain their movements with consistent deviation from the ideal trajectory.

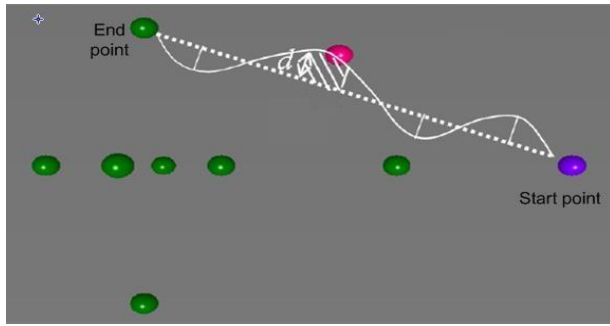


Figure 2 Deviation errors for trial trajectory

Economy of movement was computed by dividing the actual path length with the ideal path length. Actual path length was the summation of the length of straight line joining the points along the subject's trajectories. The ideal path length was measured by calculating the Euclidean distance from the starting point to the target point. The result was represented as ratio and indicated their path length's efficiency. Lower ratio values indicated the most economic path. In other words, shorter path length from initial point to a target point showed that the subjects were able to minimize their movement.

Motion smoothness was measured based on the number of zero crossings in acceleration profile. The change of velocity over time can detect the unsmooth motion. Before the velocity and acceleration profile was calculated, the displacement data was filtered by using Butterworth low pass filter at 25 Hz. This was to ensure that the high frequency noise was removed, which can produce many extra oscillations in velocity and acceleration profile. Higher number of zero crossing in acceleration profile indicates unsmooth trajectories because change in acceleration meant sudden unpredictable jerk was detected.

End-point accuracy was measured by averaging the Euclidean distance between the cursor point from the target point on the trial. The errors from x, y, z components were combined to get the resultant errors which represented end-point accuracy error. Lower error values represented higher accuracy.

End-point precision was measured using the standard deviation of the Euclidean distance between the cursor movements and the target point on the trial. It measured how well the subjects consistently maintained their hand positions at the same place. Lower values represented higher precision.

The results for all parameters from subjects were gathered, normalised and separated into two groups, the surgeon group and the non-surgeon group. The average values for each parameter were calculated to give the overall performance of the two groups for comparison. Next, for each parameter and each group, data was further divided into three

categories based on different target locations, which varied at different horizontal, vertical and depth location, to analyse how each of the target location influenced the subject's movement. Statistical analysis was performed using SPSS software. On initial analysis, all data were not normally distributed. Hence a non-parametric test was used to identify the significant difference in performance parameters between two groups. The non-parametric test chosen for this analysis was the Mann-Whitney U test.

3.0 RESULTS

Accuracy of surgeon's movement was higher and surgeons performed better with less deviation error compared to non-surgeon group. The average normalised deviation error for the surgeons was 0.13 ± 0.12 and non-surgeon, 0.17 ± 0.12 ($p < 0.05$). Figure 3 shows the results for motion path accuracy when analysed based on separate target locations. The result shows that the normalised errors for surgeons and non-surgeons were 0.152 ± 0.131 and 0.179 ± 0.119 (for horizontal targets), 0.158 ± 0.162 and 0.200 ± 0.120 (for vertical targets) and 0.223 ± 0.160 and 0.359 ± 0.245 (for depth targets) respectively. The difference between surgeons and non-surgeons was statistically significant ($p < 0.05$) for the horizontal, vertical and depth targets.

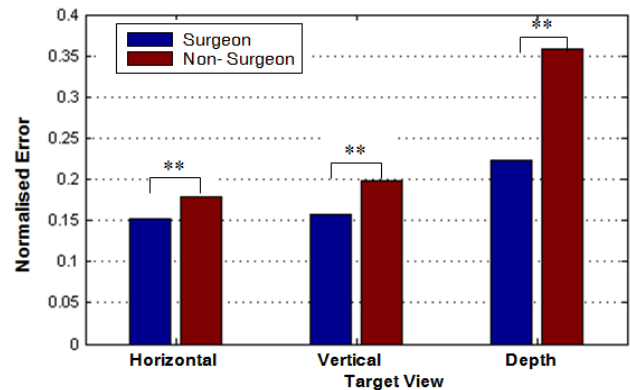


Figure 3 Motion path accuracy showing the normalised error for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

For motion path precision, surgeons made more precise movements as compared to non-surgeons. The standard deviation of deviation error was smaller in the surgeon group compared to non-surgeon group with the average value of 0.08 ± 0.11 and 0.10 ± 0.10 respectively ($p < 0.05$). When analysed based on different target locations (Figure 4), the normalised errors for surgeons and non-surgeons were 0.093 ± 0.130 and 0.093 ± 0.094 (for horizontal targets), 0.107 ± 0.165 and 0.128 ± 0.103 (for vertical targets) and 0.137 ± 0.141 and 0.282 ± 0.257 (for depth targets) respectively. The difference between surgeons and non-surgeons was statistically

significant ($p < 0.05$) for the vertical and depth targets, but it did not reach statistical significance for the horizontal target.

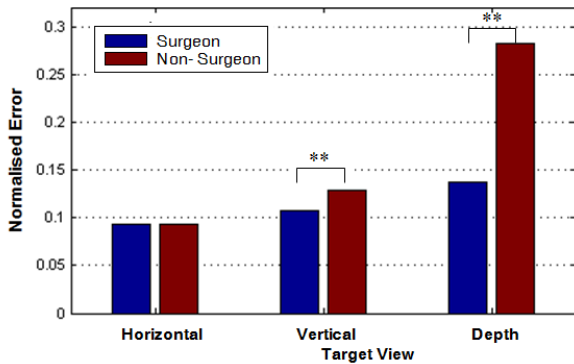


Figure 4 Motion path precision showing the normalised error for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

The surgeon group showed a lower ratio of actual path length over ideal path length, which indicated more economical movements when compared to non-surgeon's group with the average value of 3.76 ± 1.76 and 4.08 ± 2.24 respectively ($p < 0.05$). When comparing based on different target locations (Figure 5), the ratio value for surgeons and non-surgeon were 3.94 ± 2.15 and 4.23 ± 2.14 (for horizontal targets), 3.55 ± 1.58 and 4.27 ± 3.07 (for vertical targets) and 3.70 ± 1.19 and 3.70 ± 1.06 (for depth targets) respectively. The difference was statistically significant at horizontal and vertical test ($p < 0.05$). However, no significant difference was detected when the position targets were varied based on depth location ($p > 0.05$).

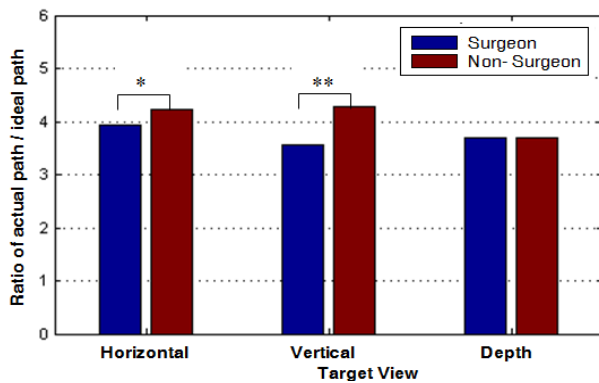


Figure 5 Economy movement showing the ratio of actual path over ideal path for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

When comparing the number of zero crossings across the two groups, the average values for surgeon and non-surgeon groups were 0.29 ± 0.14 and 0.28 ± 0.15 respectively. However, this difference was

not statistically significant indicating that surgeons and non-surgeons had similar smoothness in their trajectories when performing the reaching task using their dominant hand. Normalised value for surgeons and non-surgeons were 0.29 ± 0.12 and 0.28 ± 0.14 (for horizontal targets), 0.21 ± 0.09 and 0.24 ± 0.16 (for vertical targets) and 0.35 ± 0.18 and 0.33 ± 0.13 (for depth targets) respectively and the difference was insignificant for all target locations (Figure 6).

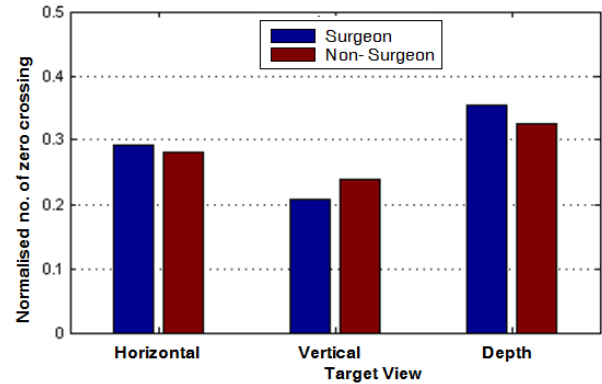


Figure 6 Motion smoothness showing the normalised number of zero crossing for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

The surgeon group was able to move their hands more accurately to their target positions compared to the other group. End-point accuracy results show that the surgeon group had lower mean deviation error compared to the non-surgeon group with the average value of 0.12 ± 0.10 and 0.17 ± 0.11 respectively ($p < 0.05$). Based on Figure 7, normalised errors for surgeons and non-surgeons were 0.129 ± 0.135 and 0.193 ± 0.128 (for horizontal targets), 0.1302 ± 0.073 and 0.184 ± 0.113 (for vertical targets) and 0.204 ± 0.117 and 0.253 ± 0.1314 (for depth targets) respectively. The Mann-Whitney result also revealed that the difference was statistically significant for all target locations.

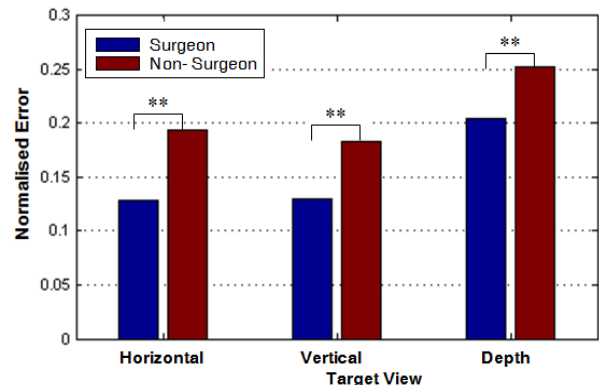


Figure 7 End-point accuracy showing the normalised error for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

A similar trend was found in end-point precision where the normalised standard deviation error was smaller for surgeon group compared to non-surgeon group. The average values were 0.04 ± 0.10 and 0.07 ± 0.10 for surgeons and non-surgeons respectively ($p < 0.05$). Surgeons were able to maintain their hand positions more steadily and precisely compared to non-surgeons. When comparing based on different target locations, the average values for surgeons and non-surgeons were 0.056 ± 0.148 and 0.088 ± 0.135 (for horizontal targets), 0.035 ± 0.035 and 0.058 ± 0.091 (for vertical targets) and 0.124 ± 0.105 and 0.156 ± 0.123 (for depth targets) respectively (Figure 8). The difference was statistically significant for all target positions at varying horizontal, vertical, and depth location with $p < 0.05$.

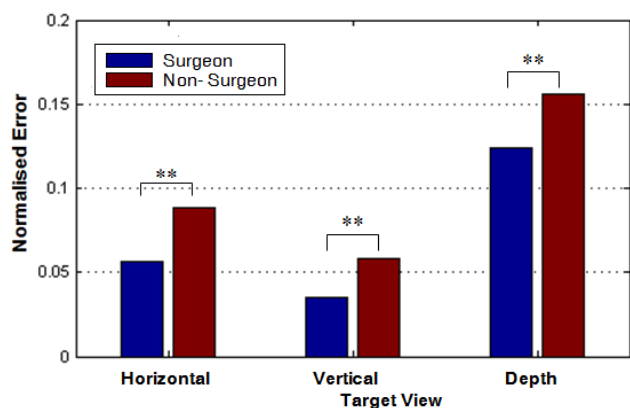


Figure 8 End-point precision showing the normalised error for surgeons (blue) and non-surgeons (red) when the targets are varied at different horizontal, vertical, and depth location

4.0 DISCUSSION

Six parameters were identified and used to compare the performance between surgeon and non-surgeon groups. An interesting observation from this experiment is that subject's performance was clearly affected by target locations. Experimental results from motion path accuracy and precision showed that mean difference in errors between the two groups were larger for targets located at different depth plane from the starting point, as compared to targets at different horizontal or vertical locations. The non-surgeon group recorded more errors when targets varied in depth and vertical positions compared to the targets that varied in horizontal position. For the surgeon group, the errors found at varying depth and horizontal positions were almost similar. At the horizontal and vertical plane, subjects could easily use their visual information to correct the errors. However, when target position varied in depth, the correct execution becomes difficult due to depth perception. In previous study, Su *et al.* conducted an experiment to investigate micromanipulation learning by using divergent force and found that the errors in z-direction (depth) for

both control and test groups were higher compared to y-direction error, indicating that the amount of error produced is affected by depth perception [24]. With more experience in real life surgical procedures, surgeons have better control in their eye-hand coordination with limited depth perception. This is probably because surgeons are more adapted to depth perception due to their exposure with microscope usage [25-26]. In addition, surgeons automate for most psychomotor skill and visual spatial perception, which has been considered as essential surgical skill [27]. Our results are also consistent with the findings of previous study [11], where Chan *et al.* showed that expert received higher mean rating score with the range of mean score from 4 to 4.5 compared to trainees who received the mean score range about 2 to 3.5 on visual spatial performance when measured using structured human grading.

One factor that affects hand dexterity is tremor. A very common tremor in normal human is physiologic tremor [28]. It is an uncontrolled movement that is inherent in all human motions [29] and age is expected to cause greater tremors in human hand [30]. Deviation errors caused by tremor would have large implications, especially in microsurgery procedures. Our study showed that the older surgeons were good in controlling their hand steadiness with lower errors in end-point accuracy and precision compared to the younger non-surgeon subjects. This is likely because surgeons reduce their tremor through slow breathing and muscle control.

Even though it is expected that the surgeons have smoother motion and would produce lower number of zero crossing compared to non-surgeon group, the results of this study showed no statistically significant difference between the two groups. The result is consistent with studies when performing laparoscopy, where the motion smoothness parameter did not prove to be significantly different between the groups [31-33]. This could be related to the experimental design, where the task might be very easy for both groups compared to real surgical procedures, hence, difference in motion smoothness was hard to capture. In addition, the setup for this task used only one hand, thus producing lower proprioceptive information which may affect the motion smoothness.

5.0 CONCLUSION

Development of assessment tools using computer based measurements can provide quantitative, measurable assessment indicators to assess psychomotor skills. These objective measurements can complement the current rating-based assessment method during the course of the training. Trainee can identify their mistakes immediately and correct their performances, without depending only on the availability of expert observations. In this study, the Green Target module was developed

using virtual reality system to compare the performances between surgeon and non-surgeon group based on a simple reaching and pointing task. The experiment showed that surgeon performed better than the non-surgeon group, with statistically significant differences for almost all parameters investigated. The parameters were motion path accuracy, motion path precision, economy of movement, end-point accuracy and end-point precision. These findings provide useful information in assessing the performance of basic surgical skill using more objective measurements. In future, more experiments can be conducted with bigger subject population and using different types of experimental design, such as using bimanual settings, where both hands can be involved during a task.

Acknowledgement

This work was supported by Ministry of Higher Education Malaysia and Universiti Teknologi Malaysia Institutional Research Grant Q.J130000.2523.12H34. We thank the subjects for their participation in this study.

References

- [1] Lossing, A. G., Hatswell, E. M., Gilas, T., Reznick, R. K. and Smith, L. C. 1992. A Technical-Skills Course for 1st-Year Residents in General Surgery: Residents in General Surgery: A Descriptive Study. *Canadian Journal of Surgery*. 35(5): 536-540.
- [2] Rodriguez-Paz, J. M., Kennedy, M., Salas, E., Wu, A. W., Sexton, J. B., Hunt, E. A. and Pronovost, P. J. 2009. Beyond 'See One, Do One, Teach One': Toward a Different Training Paradigm. *Postgraduate Medical Journal*. 85(1003): 244-249.
- [3] Bridges, M. and Diamond, D. L. 1999. The Financial Impact of Teaching Surgical Residents in the Operating Room. *American Journal of Surgery*. 177(1): 28-32.
- [4] Babineau, T. J., Becker, J., Gibbons, G., Sentovic, S., Hess, D., Robertson, S. and Stone, M. 2004. The 'Cost' of Operative Training for Surgical Residents. *Archives of Surgery*. 139(4): 366-369.
- [5] Moorthy, K., Munz, Y., Adams, S., Pandey, V. and Darzi, A. 2005. A Human Factors Analysis of Technical and Team Skills Among Surgical Trainees During Procedural Simulations in a Simulated Operating Theatre. *Annals of Surgery*. 242(5): 631-639.
- [6] Galasko, C. and Mackay, C. 1997. Unsupervised Surgical Training. Logbooks Are Essential for Assessing Progress. *British Medical Journal*. 315(7118): 1306-1307.
- [7] Begg, C. B., Cramer, L. D., Hoskins, W. J. and Brennan, M. F. 1998. Impact of Hospital Volume on Operative Mortality for Major Cancer Surgery. *Journal of the American Medical Association*. 280(20): 1747-1751.
- [8] Bridgewater, B., Grayson, A. D., Jackson, M., Brooks, N., Grotte, G. J., Keenan, D. J. M., Millner, R., Fabri, B. M. and Jones, M. 2003. Surgeon Specific Mortality in Adult Cardiac Surgery: Comparison Between Crude and Risk Stratified Data. *British Medical Journal*. 327(7405): 13-17.
- [9] Turrentine, F. E., Wang, H., Simpson, V. B. and Jones, R. S. 2006. Surgical Risk Factors, Morbidity, and Mortality in Elderly Patients. *Journal of the American College of Surgeons*. 203(6): 865-877.
- [10] Doyle, J. D., Webber, E. M. and Sidhu, R. S. 2007. A Universal Global Rating Scale for the Evaluation of Technical Skills in the Operating Room. *American Journal of Surgery*. 193(5): 551-555.
- [11] Chan, W., Niranjana, N. and Ramakrishnan, V. 2010. Structured Assessment of Microsurgery Skills in the Clinical Setting. *Journal of Plastic, Reconstructive & Aesthetic Surgery*. 63(8): 1329-1334.
- [12] Kalu, P. U., Atkins, J., Baker, D., Green, C. J. and Butler, P. E. M. 2005. How Do We Assess Microsurgical Skill?. *Microsurgery*. 25(1): 25-29.
- [13] Bann, S., Davis, I. M., Moorthy, K., Munz, Y., Hernandez, J., Khan, M., Datta, V. and Darzi, A. 2005. The Reliability of Multiple Objective Measures of Surgery and the Role of Human Performance. *American Journal of Surgery*. 189(6): 747-52.
- [14] Goff, B. A., Lentz, G. M., Lee, D., Fenner, D., Morris, J. and Mandel, L. S. 2001. Development of A Bench Station Objective Structured Assessment of Technical Skills. *Obstetrics and Gynecology*. 98(3): 412-416.
- [15] Wanzel, K. R., Matsumoto, E. D., Hamstra, S. J. and Anastakis, D. J. 2002. Teaching Technical Skills: Training on a Simple, Inexpensive, and Portable Model. *Journal of Plastic, Reconstructive & Aesthetic Surgery*. 109(1): 258-263.
- [16] Grober, E. D., Hamstra, S. J., Wanzel, K. R., Reznick, R. K. and Matsumoto, E. D. 2003. Validation of Novel and Objective Measures of Microsurgical Skill: Hand-Motion Analysis and Stereoscopic Visual Acuity. *Microsurgery*. 23(4): 317-322.
- [17] Bann, S. D., Khan, M. S. and Darzi, A. W. 2003. Measurement of Surgical Dexterity Using Motion Analysis of Simple Bench Tasks. *World Journal of Surgery*. 27(4): 390-394.
- [18] Blyth, P., Stott, N. S. and Anderson, I. A. 2008. Virtual Reality Assessment of Technical Skill Using the Bonedoc DHS Simulator. *Injury, International Journal of the Care of The Injured*. 39(10): 1127-1133.
- [19] Then, T. F., Su, E. L. M., Ahmmad, S. N. Z. and Yeong, C. F. 2013. Massed Training Versus Interval Training for Computer-Based Suturing Skill Acquisition. *Journal of Medical Imaging and Health Informatics*. 3(4): 503-508.
- [20] Ahmmad, S. N. Z., Su, E. L. M. and Yeong, C. F. 2013. Effect of Intermittent Haptic Disturbance in Motor Skill Training. *Applied Mechanics and Materials*. 432: 403-408.
- [21] Pellen, M. G. C., Horgan, L. F., Barton, J. R. and Attwood, S. E. 2009. Construct Validity of the Promis Laparoscopic Simulator. *Surgical Endoscopy*. 23(1): 130-139.
- [22] Ahmmad, S. N. Z., Chew, Z. S., Su, E. L. M. and Yeong, C. F. 2015. Force Variability as an Objective Measure of Surgical Skill. *Jurnal Teknologi*. 74(6): 125-128.
- [23] Ahmmad, S. N. Z., Su, E. L. M., Yeong, C. F. and Narayanan, A. L. T. 2014. Experimental Study of Surgeon's Psychomotor Skill Using Sensor-Based Measurement. *Procedia Computer Science*. 42: 130-137.
- [24] Su, E. L. M., Ganesh, G., Yeong, C. F. and Burdet, E. 2010. Accurate Micromanipulation Induced by Performing in Unstable Dynamics. *19th IEEE International Symposium on Robot and Human Interactive Communication*. 762-766.
- [25] Harwell, R. C. & Ferguson, R. L. 1983. Physiologic Tremor and Microsurgery. *Microsurgery*. 4(3): 187-192.
- [26] Rooks, M. D., Slappey, J. and Zusanis, K. 1993. Precision of Suture Placement With Microscope- And Loupe-Assisted Anastomoses. *Microsurgery*. 14(8): 547-550.
- [27] Risucci, D. A. 2002. Visual Spatial Perception and Surgical Competence. *American Journal of surgery*. 184(3): 291-295.
- [28] Elble, R. J. 2003. Characteristics of Physiologic Tremor in Young and Elderly Adults. *Journal of Clinical Neurophysiology*. 114(4): 624-635.
- [29] Elble, R. J. 1996. Central Mechanisms of Tremor. *Journal of Clinical Neurophysiology*. 13(2): 133-144.

- [30] Sturman, M. M., Vaillancourt, D. E. and Corcos, D. M. 2005. Effects of Aging on The Regularity of Physiological Tremor. *Journal of Neurophysiology*. 93(6): 3064-3074.
- [31] Hofstad, E. F., Vapenstad, C., Chmarra, M. K., Lango, T., Kuhry, E. and Marvik, R. 2012. A Study of Psychomotor Skills in Minimally Invasive Surgery: What Differentiates Expert and Nonexpert Performance. *Surgical Endoscopy*. 27(3): 854-863.
- [32] Chmarra, M. K., Kolkman, W., Jansen, F. W., Grimbergen, C. A. and Dankelman, J. 2007. The Influence of Experience and Camera Holding on Laparoscopic Instrument Movements Measured With the Trendero Tracking System. *Surgical Endoscopy*. 21(11): 2069-2075.
- [33] Maithel, S. K., Villegas, L., Stylopoulos, N., Dawson, S. and Jones, D. B. 2005. Simulated Laparoscopy Using a Head-Mounted Display vs Traditional Video Monitor: An Assessment of Performance and Muscle Fatigue. *Surgical Endoscopy*. 19: 406-411.