

## Objective assessment of surgeon's psychomotor skill using virtual reality module

Siti Nor Zawani Ahmmad<sup>1</sup>, Eileen Su Lee Ming<sup>2</sup>, Yeong Che Fai<sup>3</sup>, Suneet Sood<sup>4</sup>, Anil Gandhi<sup>5</sup>,  
Nur Syarafina Mohamed<sup>6</sup>, Hisyam Abdul Rahman<sup>7</sup>, Etienne Burdet<sup>8</sup>

<sup>1,6</sup>Universiti Kuala Lumpur, Malaysian Institute of Industrial Technology, Malaysia

<sup>2</sup>School of Electrical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, Malaysia

<sup>3</sup>Centre of Artificial Intelligence and Robotics (CAIRO), Universiti Teknologi Malaysia, Malaysia

<sup>4,5</sup>Department of Surgery, Faculty of Medicine, Jeffrey Cheah School of Medicine and Health Sciences,  
Monash University, Malaysia

<sup>7</sup>Department of Robotic and Mechatronic Engineering, Faculty of Electrical and Electronic Engineering,  
Universiti Tun Hussein Onn Malaysia, Malaysia

<sup>8</sup>Department of Bioengineering, Imperial College London, South Kensington Campus, United Kingdom

---

---

### Article Info

#### Article history:

Received Dec 13, 2018

Revised Feb 14, 2019

Accepted Feb 27, 2019

---

#### Keywords:

Assessment parameter

Basic surgical skill

Psychomotor skill

Virtual reality

---

---

### ABSTRACT

This study aims to identify measurable parameters that could be used as objective assessment parameters to evaluate surgical dexterity using computer-based assessment module. A virtual reality module was developed to measure dynamic and static hand movements in a bimanual experimental setting. The experiment was conducted with sixteen subjects divided into two groups: surgeons (N = 5) and non-surgeons (N = 11). Results showed that surgeons outperformed the non-surgeons in motion path accuracy, motion path precision, economy of movement, motion smoothness, end-point accuracy and end-point precision. The six objective parameters can complement existing assessment methods to better quantify a trainee's performance. These parameters also could provide information of hand movements that cannot be measured with the human eye. An assessment strategy using appropriate parameters could help trainees learn on computer-based systems, identify their mistakes and improve their skill towards the competency, without relying too much on bench models and cadavers.

*Copyright © 2019 Institute of Advanced Engineering and Science.  
All rights reserved.*

---

### Corresponding Author:

Siti Nor Zawani Ahmmad,  
Universiti Kuala Lumpur,  
Malaysian Institute of Industrial Technology, Malaysia.  
Email: sitinorzawani@unikl.edu.my

---

## 1. INTRODUCTION

A high level of psychomotor skill is required to perform a surgical procedure safely and dexterously. Complicated operations require much training and practice, especially for microsurgical procedures, and microsurgery skill is affected by many factors, such as perceptual and motor limitations that steepen the learning curve [1]. This has led to the development of training strategies in the laboratory, where trainees need to achieve a certain competency level before operating on patients [2]-[3].

Studies related with skill acquisition in various surgical fields have shown that certain "aptitudes", such as visual-spatial and psychomotor skill, are strongly connected with the ability to acquire surgical skill [1]. Different individuals have different aptitudes, where these can be considered as important for the development of skill in specific areas of surgery. In previous studies on minimally invasive surgery, visual-spatial and psychomotor ability have been associated with performance and learning [4]-[6]. Similarly, in the field of microsurgery, depth and visual-spatial awareness are important for trainees to manipulate tools efficiently under the microscope. In 1994, Murdoch et al. [7] found significant correlation

between visual spatial and psychomotor skill and microsurgical skill. In 2012, Nugent et al. [1] assessed the performance of trainee on day 2 and 5 during microsurgical training using Objective Structured Assessment of Technical Skills (OSATS). They found that participants' skills were significantly improved, and the degree of skill improvement strongly correlated with psychomotor aptitude assessment score.

The increasingly sophisticated technologies are able to provide opportunities for surgeons to improve their surgical performance through various training programs. Equally important to learning skill is the assessment method, as it determines the learning and technical proficiency of the trainees.

## 2. LITERATURE REVIEW

Traditional assessment method is based on direct observation and human grading. The expert surgeon typically observe the trainee performance and provide verbal feedback. This assessment is subjective, and can be unreliable as it is affected by inter and intra-rater variability and recall bias [8]. There are few types of successful formative assessments used in surgical skill which are based on human grading. The Objective Structured Clinical Examination (OSCE) is one of the assessment designed for the laboratory setting used to assess clinical performance of trainees at various clinical stations using specific checklist [9]. Candidate rotates through all the station to perform the task given and the examiners present at each station will give marks in standardised marking process. OSCE stations involve several tasks including essays tests and clinical situation, where the student interacts with the patient. It mostly focuses on the assessment of procedural knowledge on clinical scenario, interpretation and discussion of diagnosis, behaviour and attitude of the trainee towards the patient. However, technical evaluation of psychomotor skills is not really given attention [10].

Another type of formative assessment is the Objective Structured Assessment of Technical Skills (OSATS). This is most commonly used for assessment in open and laparoscopy surgery. It was first reported by Martin *et al.* [11] and consisted of checklist and global rating scale assessed with five point anchored Likert scale [12]. Candidate performed a standardised operative task under direct observation from expert. OSATS was originally designed and validated using bench model (inanimate) and has been extensively validated in general surgery [13], [14]. Since then, the concept of global rating scale had been modified by numerous researchers in other surgical approaches and other types of setting [15], [16]. The drawback of the system is that although the rating scale and criteria have been well defined, expert raters may still be influenced by the halo effect and bias based on the background level of trainee.

Many researchers have been eager to find objective and quantitative methods for assessment of surgical skills using computational based system such as simulator [17], [18]. Currently, most simulators that assess basic psychomotor skill are available for minimally invasive surgery: examples include MIST-VR, LapMentor, LapSim and ProMIS [19]. In contrast, very few simulators assess skills used in microsurgery [20]-[22]. More importantly, effective assessment parameters to measure basic microsurgical skill have not been clearly identified, and are still under study [23]-[25]. Hand motion analysis, such as Imperial College Surgical Assessment Device (ICSAD), is a device proposed for use in objective assessment of microsurgical skill [26]-[28]. However, measurement parameters are limited. Only a few parameters can be identified - time, total number of hand movements and total hand travel distance. Combining additional and different parameters may provide a better way of measuring trainees' performances.

Virtual reality have been widely used in many application such as in teleoperation system, geographic information and practical teaching [29]-[31]. In this study, a virtual reality assessment module based on reaching and pointing tasks has been designed to analyze hand movement data. This study also proposes a bimanual setting where the dominant hand is used to perform tasks, while the non-dominant hand provides proprioceptive and visual reference for the target position. This task mimics real microsurgery procedures, where a surgeon sometimes moves the tool out of the viewing zone, and then brings the tool back to the operative field. The presence of the left hand at the target position could perhaps help the right hand to move accurately toward the target position when the visual feedback is missing. A typical surgical procedure involves both hands and this bimanual setting will investigate how subject's performance could be affected by the use of the non-dominant hand. In addition, the viewing methodology is designed closely similar with the view in microsurgical procedure, which is using stereo view.

The aim of this study is to extract a set of measurable parameters from movement data to determine the parameters that can be used to differentiate the performance between surgeon and non-surgeon groups. These measurement parameters can be used in the construction and validation of an assessment tool to provide feedback to users undergoing technical skill development. Furthermore, measuring individual baseline skills in the early phase of training could help to speed up skill acquisition, highlighting what is good and what can be improved in trainee's performance.

### 3. RESEARCH METHOD

Sixteen subjects without known impairment participated in this study. They were divided into two groups: surgeon and non-surgeon. The 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. The non-surgeon's group comprised eleven healthy adults (female=5, male=6) aged between 22 and 26 years' old. All subjects declared to be right-hand dominant. The details of the nature and purpose of the research were explained to them, and consent was obtained for participation in the study.

#### 3.1. Experimental Setup

The PHANTOM Omni haptic device from Sensable Technologies was used for position measurement during movement. The haptic feedback loop ran at 1000 Hz. This haptic device provides 6 degree of freedom (DOF) positional and orientation sensing using digital encoders with nominal accuracy of 0.055 mm. In addition, 3-DOF force feedback can be provided with a continuous force of 0.88 N and maximum force of 3.3 N within  $160 \times 120 \times 70\text{mm}^3$ .

The basic user interface framework for the task module was developed using Microsoft Visual Studio C++. The graphics of virtual environment and objects were 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 3-D monitor Acer HS244HQ with a pair of active 3-D shutter glasses (built-in IR emitter). The 3-D monitor has a 23.6 inch display with  $1920 \times 1080$  pixel full HD resolution at 120Hz refresh rate. The graphics card used with this study was the nVidia GeForce 54m series.

An experimental software module was developed to investigate subject movement. During the experiment, subjects were able to perceive their movements in three-dimensions, and could visually estimate the depth of their movements using a pair of shutter glasses and the 3-D feature of the monitor screen. A visual magnification of  $10\times$  was used to amplify movement error throughout the trial. 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 positive pointed upwards while the z-axis positive pointed towards the viewer. The visual display on the screen consisted of the starting point shown as a static purple sphere inside a yellow box. A pink sphere acted as the movement of the dominant hand and an orange sphere acted as the movement of non-dominant hand. In the middle region there were 7 green spheres with different co-ordinates which represented the targets for the subject to aim at. These targets were presented one at a time is shown in Figure 1 (a).

Each subject performed a set of bimanual procedures using two PHANTOM Omni haptic devices, one for the dominant hand and one for the non-dominant hand. All subjects were shown a short video before commencing the experimental trials. This video showed them how the green target module worked, and subjects were given time to practice. Once the subjects understood the procedure, they proceeded to start the experiment. At the start of the session, each subject was required to be seated comfortably at the table with the eyes levelled to the middle of the screen, looking at the 3-D monitor screen through the shutter glasses is shown in Figure 1 (b). The PHANTOM was placed 20 cm away from the subject's torso. 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, a key was pressed to set the pink sphere to overlap with the starting point, and the first target appeared in the form of a green sphere. Simultaneously, the non-dominant hand, which was represented as an orange sphere, would be moved automatically by haptic forces and positioned near to the green target. 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. The left hand was held in place with haptic forces from the beginning of the trial until the subject successfully finished the trial. When the next trial was about to begin, the left stylus would be moved again towards the next target and the right stylus would be moved to the starting point. These processes were repeated for all the targets. 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 ( $7 \text{ targets} \times 3 \text{ repetitions} \times 2 \text{ sessions}$ ).

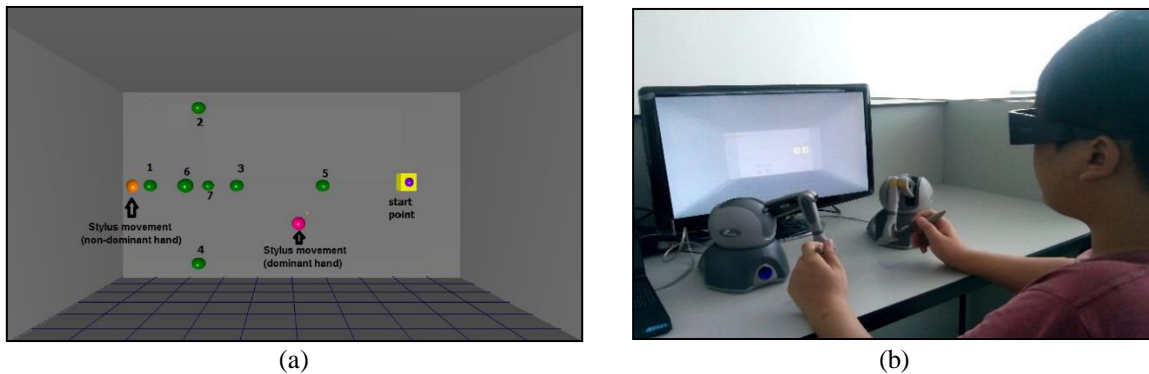


Figure 1. (a). Visual display on the 3-D monitor screen consists of starting point showed as static purple sphere inside a yellow box, a pink sphere acted as the cursor for right hand movement, a green sphere acted as the target and orange sphere acted as the left hand movement. (b) Experimental set-up for bimanual setting. Subject held two haptics stylus, one in each hand and wearing shutter glasses to perform the experiment

### 3.2. Data Analysis

Data collected was processed using MATLAB (The Mathworks, USA). Statistical analysis was performed using Statistical Packages for the Social Sciences (SPSS) version 16. From movement trajectories data, several useful parameters were extracted and processed.

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 analyze dynamic movements while the pointing data was used to analyze static movement of subjects. The extraction of data for reaching analysis started from the moment subjects moved their cursor until they reached the target point. For pointing analysis, data was extracted within a three-second time frame when subjects reached the target and held the cursor for at least 3s within 0.5 mm radius from the target distance.

The parameters analyzed in reaching analysis were motion path accuracy and precision, economy of movement and motion smoothness. For pointing analysis, two parameters were analyzed: end-point accuracy and end-point precision.

Motion path accuracy was identified by calculating the average distance,  $d$  throughout the trajectory trial relative to an ideal straight line trajectory as shown in Figure 2. Outlier data was removed before calculating the mean of  $d$ . The outliers are detected using the Thompson's Tau method and used a two-sided test with a significance level of  $\alpha = 0.05$  for the outlier removal process. A lower  $d$  indicated higher motion path accuracy.

Motion path precision was identified by calculating the standard deviation of the deviation errors,  $d$ , throughout the trajectory trial. The deviation error in each discrete point of trial was gathered and the outlier data was removed before calculating the standard deviation of  $d$ . Smaller values indicated higher precision.

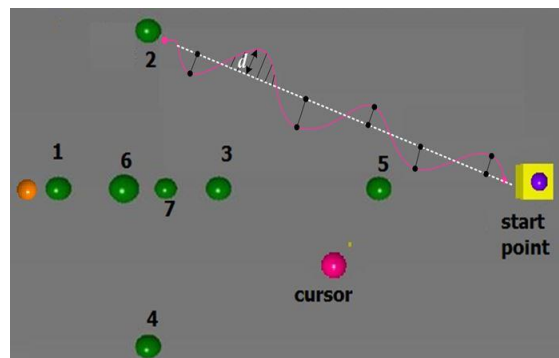


Figure 2. Deviation error,  $d$  was represented as the path error made by the subjects based on an ideal straight line trajectory (indicated by straight white dotted line)

Economy of movement was computed by dividing the actual path length with the ideal path length. Actual path length was measured by summation of the length of straight lines 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 a ratio and indicated their path length's efficiency. Lower ratios indicated the most economic paths. Shorter path lengths from initial points to target points showed that the subjects were able to minimize their movement. The path length was measured based on (1):

$$\text{Path length} = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \tag{1}$$

where

- $i$  :  $i^{\text{th}}$  sample in discrete data stream,
- $N$  : total number of data points in selected motion path
- $x_i, y_i, z_i$  : instantaneous coordinate points at  $i^{\text{th}}$  sample in  $x, y,$  and  $z$  direction

*Motion smoothness* was measured based on the number of zero crossings in the acceleration profile [32]. The change of velocity over time can detect unsmooth motion. Before the velocity and acceleration were computed numerically, the displacement data was filtered using a Butterworth low pass filter at 25 Hz. This was to ensure removal of the high frequency noise, which can produce many extra oscillations in velocity and acceleration profile. Higher numbers of zero crossing in the acceleration profile indicated unsmooth trajectories, because change in acceleration meant that a sudden unpredictable jerk was detected.

*End-point accuracy* was measured by averaging the Euclidean distance between the pointing data movement from the target point. Initially, the errors from the  $x, y,$  and  $z$  components were combined to get the resultant errors, as shown in (2). Then, the error in each discrete point of the trial was gathered, and the outlier data was removed before calculating the mean of distance error, which was represented as end-point accuracy error. Lower error values represented higher accuracy.

$$\text{Resultant Error} = \sum_{i=1}^N \sqrt{(x_i - x_t)^2 + (y_i - y_t)^2 + (z_i - z_t)^2} \tag{2}$$

where

- $i$  :  $i^{\text{th}}$  sample in discrete data stream,
- $N$  : total number of data points in selected motion path
- $x_i, y_i, z_i$  : instantaneous coordinate points at  $i^{\text{th}}$  sample in  $x, y,$  and  $z$  direction
- $x_t, y_t, z_t$  : target coordinate point in  $x, y,$  and  $z$  direction

*End-point precision* was measured using the standard deviation of the Euclidean distance between the pointing data movement and the target point on the trial. The distance error in each discrete point of trial was gathered and the outlier data were removed before calculating the standard deviation of distance error. This parameter indicated 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, normalized 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 for the comparison. Then, all subjects' data were divided into three categories depending on the target characteristic of experimental data to analyze how each of the target locations (horizontal, vertical and depth) influenced the subject's movement.

For horizontal targets, Target 1, Target 3 and Target 5 were averaged as one value because they shared the same plane but varied at different  $x$ - location. For vertical targets, Target 2 and Target 4 were grouped together because they shared the same plane, but had different locations on the  $y$ -axis. For depth target, Target 6 and Target 7 were grouped for analysis because they had same location in  $x$  and  $y$  directions but had different location in  $z$ - direction as shown in Figure 3.

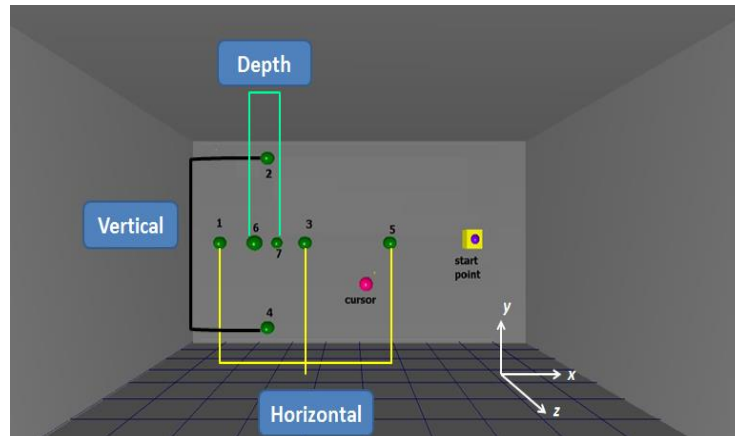


Figure 3. Target locations based on horizontal, vertical and depth position

### 3.3. Statistical Analysis

In order to identify the significant difference of performance between two groups, statistical tests were performed using SPSS software. On initial analysis, all data were tested using Shapiro-Wilk test to identify the normality of data distribution. The data was not normally distributed. Hence a non-parametric test was used to identify the significant differences in performance parameters between two groups. The non-parametric test chosen for this analysis was the Mann-Whitney U test. Figure 4 shows the summary of experimental process and data analysis for this study.

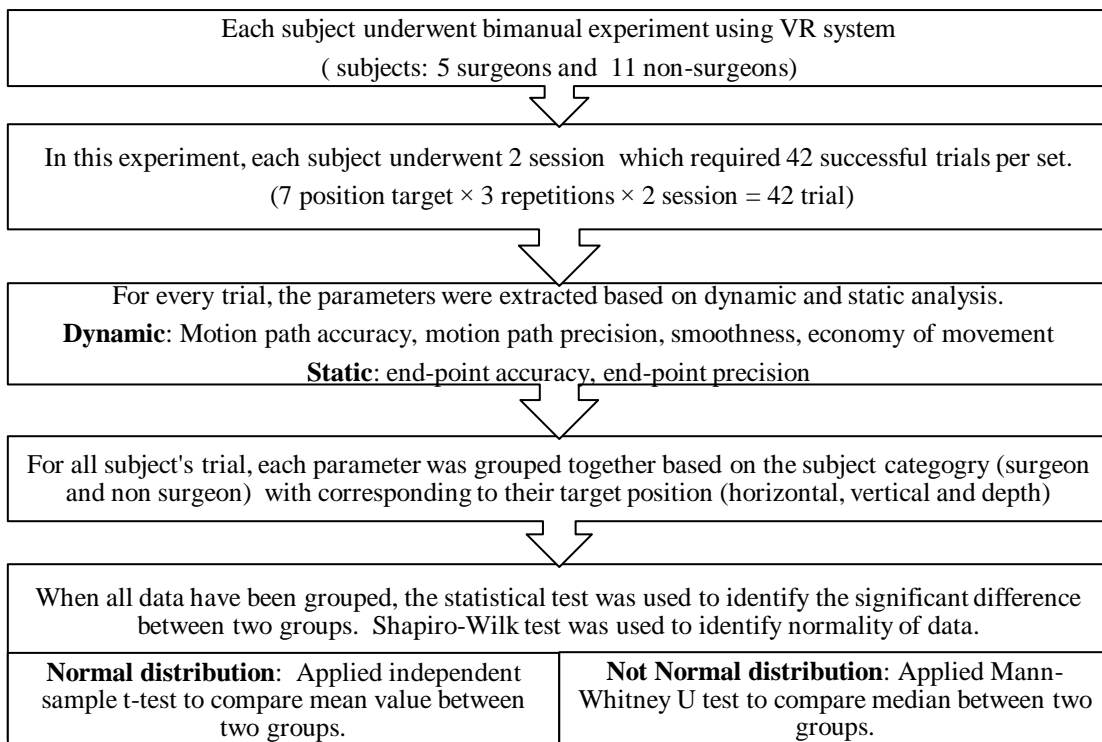


Figure 4. Summary of experimental and data analysis process

## 4. RESULTS AND ANALYSIS

In this section, it is explained the results of research and at the same time is given the comprehensive discussion.

### 4.1. Dynamic Movement Analysis

Surgeons outperformed the non-surgeons in motion path accuracy. Lower mean deviation errors for all target views were found in the surgeon group compared to non-surgeon group. The normalized error for surgeons was  $0.095 \pm 0.062$ ; and for non-surgeons it was  $0.148 \pm 0.105$ , ( $p \leq 0.001$ ). Figure 5 shows the results for motion path accuracy when analyzed based on separate target locations. The result shows that the normalized errors for surgeons and non-surgeons were  $0.092 \pm 0.060$  and  $0.114 \pm 0.865$  (for horizontal targets),  $0.168 \pm 0.114$  and  $0.273 \pm 0.166$  (for vertical targets) and  $0.1613 \pm 0.097$  and  $0.308 \pm 0.207$  (for depth targets) respectively. The performance difference between surgeons and non-surgeons was statistically significant for the horizontal ( $p = 0.01$ ), vertical ( $p \leq 0.001$ ) and depth targets ( $p \leq 0.001$ ).

In motion path precision, surgeons also performed more precise movements with lower errors 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  $0.050 \pm 0.053$  and  $0.082 \pm 0.091$  respectively ( $p \leq 0.001$ ). When analyzed based on different target locations as shown in Figure 6, the normalized errors for surgeons and non-surgeons were  $0.048 \pm 0.045$  and  $0.058 \pm 0.085$  (for horizontal targets),  $0.114 \pm 0.124$  and  $0.197 \pm 0.170$  (for vertical targets) and  $0.099 \pm 0.115$  and  $0.216 \pm 0.208$  (for depth targets) respectively. The difference between surgeons and non-surgeons was statistically significant for the vertical ( $p \leq 0.001$ ) and depth targets ( $p \leq 0.001$ ), but it did not reach statistical significance for the horizontal target ( $p = 0.12$ ).

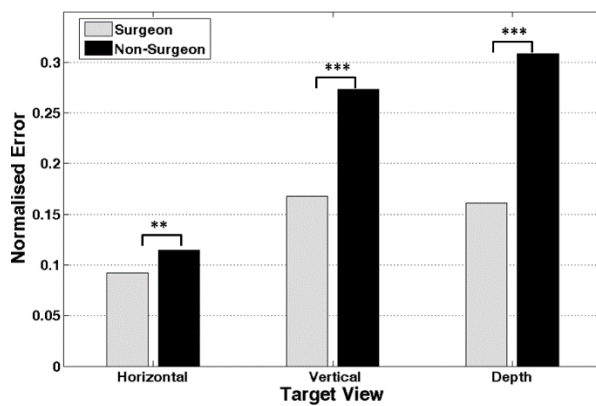


Figure 5. Motion path accuracy showing the normalized error for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\*\* indicated significance at  $p \leq 0.01$  and \*\*\* indicated significance at  $p \leq 0.001$ )

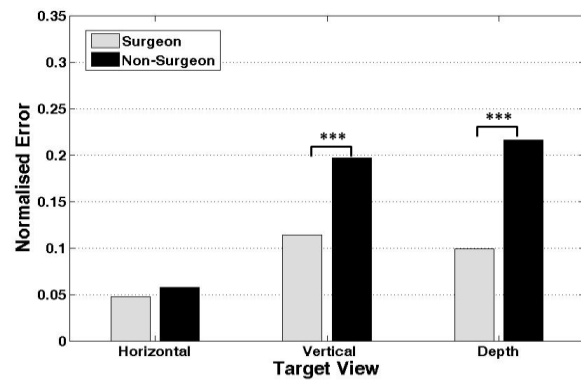


Figure 6. Motion path precision showing the normalized error for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\*\*\*) indicated significance at  $p \leq 0.001$ )

As shown in Figure 7, 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 group with the average value of  $3.229 \pm 1.248$  and  $3.671 \pm 1.352$  respectively ( $p \leq 0.001$ ). The ratio value for surgeons and non-surgeon when comparing based on difference target locations were  $3.393 \pm 1.634$  and  $3.761 \pm 1.754$  (for horizontal targets),  $3.134 \pm 0.952$  and  $3.552 \pm 0.931$  (for vertical targets) and  $3.080 \pm 0.713$  and  $3.655 \pm 0.956$  (for depth targets) respectively. The Mann-Whitney U test revealed the difference between two groups was statistically significant at horizontal ( $p=0.047$ ), vertical ( $p \leq 0.001$ ) and depth targets ( $p \leq 0.001$ ).

Normalized number of zero crossing showed lower values for surgeon group, indicating smoother movements compared to non-surgeon group. The average values were  $0.364 \pm 0.153$  and  $0.434 \pm 0.170$  respectively ( $p \leq 0.001$ ). When comparing based on different target locations, the normalized values for surgeons and non-surgeons were  $0.447 \pm 0.152$  and  $0.494 \pm 0.163$  (for horizontal targets),  $0.291 \pm 0.125$  and  $0.362 \pm 0.16$  (for vertical targets) and  $0.313 \pm 0.117$  and  $0.416 \pm 0.170$  (for depth targets) respectively is shown in Figure 8. The difference between surgeons and non-surgeons was statistically significant for horizontal ( $p=0.013$ ), vertical ( $p \leq 0.001$ ) and depth targets ( $p \leq 0.001$ ).



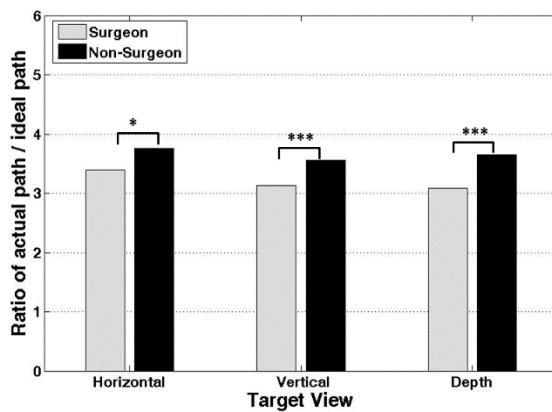


Figure 7. Economy movement showing the ratio of actual path over ideal path for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\* indicated significance at  $p \leq 0.05$  and \*\*\* indicated significance at  $p \leq 0.001$ )

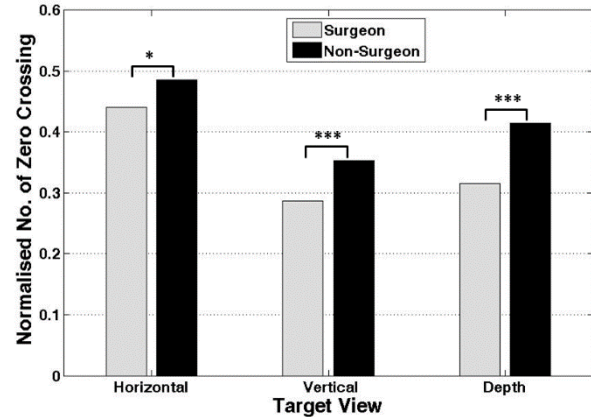


Figure 8. Motion smoothness showing the normalized number of zero crossing for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\* indicated significance at  $p \leq 0.05$  and \*\*\* indicated significance at  $p \leq 0.001$ )

#### 4.2. Static Movement Analysis

End-point accuracy result shows that the surgeon group recorded lower mean errors compared to the non-surgeon group in static movement. Surgeons were able to maintain their hand more accurately at their target position compared to non-surgeons with the average value of  $0.119 \pm 0.105$  and  $0.160 \pm 0.110$  respectively ( $p < 0.05$ ). When comparing based on different target locations (Figure 9), the normalized error for surgeons and non-surgeons were  $0.134 \pm 0.130$  and  $0.173 \pm 0.131$  (for horizontal targets),  $0.141 \pm 0.076$  and  $0.209 \pm 0.127$  (for vertical targets) and  $0.128 \pm 0.110$  and  $0.156 \pm 0.103$  (for depth targets) respectively. The difference between surgeons and non-surgeons was statistically significant for horizontal ( $p \leq 0.001$ ), vertical ( $p \leq 0.001$ ) and depth targets ( $p \leq 0.001$ ).

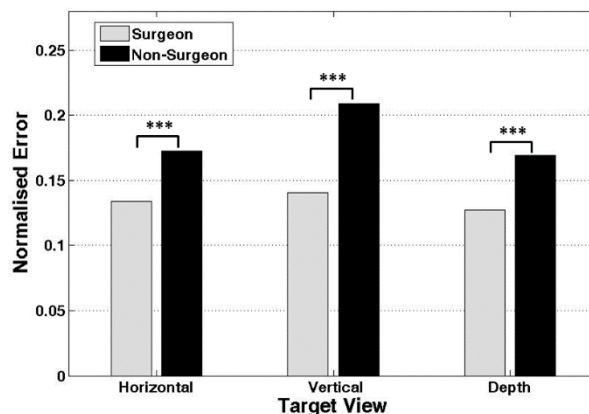


Figure 9. End-point accuracy showing the normalized error for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\*\*\*) indicated significance at  $p \leq 0.001$ )

Similar results were found in end-point precision metrics where surgeons were also able to maintain their hand position precisely compared to non-surgeons, with average values of  $0.056 \pm 0.118$  and  $0.069 \pm 0.115$  respectively,  $p \leq 0.001$ . As shown in Figure 10, the normalized error for surgeons and non-surgeons when comparing based on different target locations were  $0.079 \pm 0.152$  and  $0.091 \pm 0.147$



(for horizontal targets),  $0.036 \pm 0.032$  and  $0.068 \pm 0.118$  (for vertical targets) and  $0.054 \pm 0.130$  and  $0.059 \pm 0.061$  (for depth targets) respectively. The Mann-Whitney U test revealed the difference between two groups was statistically significant for horizontal ( $p \leq 0.001$ ), vertical ( $p \leq 0.001$ ) and depth targets ( $p = 0.01$ ).

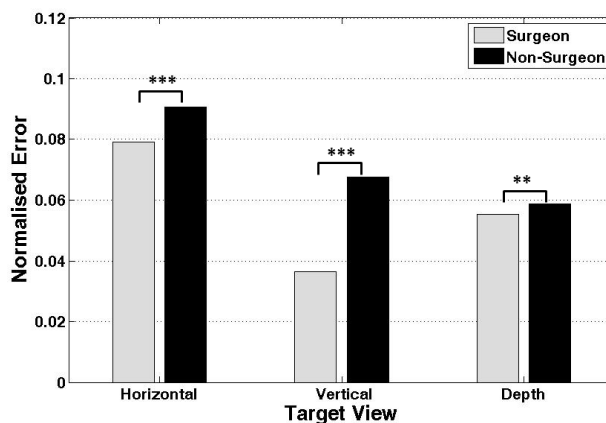


Figure 10. End-point precision showing the normalized error for surgeons (grey) and non-surgeons (black) when the targets are varied at different horizontal, vertical, and depth locations for bimanual-right experiment. (\*\* indicated significance at  $p \leq 0.01$  and \*\*\* indicated significance at  $p \leq 0.001$ ).

Surgical procedures require the coordinated movements of both hands to complete an operation. For example, during suturing task, the left hand usually holds the tissue with forceps to facilitate for needle insertion by the right hand. The left hand also holds the needle during tissue exit while waiting for the right hand to retrieve the needle again. Instead of using visual information as an input, additional proprioceptive sense from the reference left hand may help in coordinating bimanual movement [33]. A bimanual right hand setting is a better experimental setup for measuring psychomotor skills compared to our previous study using a unimanual experimental setting [34]. Using this bimanual setting, we identified six parameters which could differentiate surgeons from non-surgeons. The parameters were motion path accuracy, motion path precision, economy of movement, motion smoothness, end-point accuracy and end point precision.

A prior study on motor skill shows that the use of two hands improve proprioception [35]. In a study of psychomotor skill in minimally invasive surgery, Hofstad *et al.* [36] found that experts had significantly better bimanual dexterity compared to intermediates and novices. This report is consistent with the findings of our study, where the surgeons were able to use their skills to coordinate the haptic stylus using both hands to perform better movements in a bimanual setting compared to non-surgeons.

In dynamic analysis movement, specifically for motion path accuracy and precision (Figure 5 and Figure 6), the results showed that the mean differences in errors between the two groups were larger when targets were located at different depth locations, as compared to targets at different horizontal or vertical locations. This indicates that the subject's performance can be influenced by varying targets locations. The non-surgeon group recorded higher errors when targets varied in depth position than when targets varied in vertical and horizontal positions. This is because the correct execution becomes more difficult due to depth perception [37]. Within the surgeon group, the errors found at varying horizontal, vertical and depth positions were almost similar. In addition, skilled surgeons would automate for most psychomotor skill and visual spatial perception, which are considered as essential surgical skills [38].

Hand tremor can lead to deviation errors. In microsurgery procedures, the deviation errors can lead to high risk implications. The older the person becomes, the higher the tendency of hand tremor [30]. However, surgeons are trained to reduce tremor through slow breathing and muscle control practice. From this study, results showed that older surgeons were better in controlling hand steadiness, with lower errors in end-point accuracy and precision, compared to younger non-surgeon subjects. Despite being older than non-surgeons' group, trained surgeon still can outperformed young non-surgeon's group most likely due to tremor control training and experiences they have.

## 5. CONCLUSION

In this study, a computer-based assessment module using virtual reality system was developed to evaluate psychomotor skill for surgical training. The experiment was conducted with specific protocols and equipment for data collection. Suitable parameters for assessment module were identified and statistically analysed to determine the greatest potential for discriminating expert surgeon from non-surgeon. The findings have shown that six parameters: motion path accuracy, motion path precision, economy of movement, motion smoothness, end-point accuracy and end point precision were able to differentiate well between surgeon and non-surgeon groups. These parameters were proposed to evaluate the motor components specifically related to microsurgical skill. In addition, the proposed parameters can potentially be used to assess performance skill at any stage of motor learning. If the performance of trainees can be assessed at early stage of training, the lack of motor proficiency can be identified and interventions be given early. Next, specific training can be better tailored to meet the needs of the individual.

## REFERENCES

- [1] Nugent E, Joyce C, Perez-Abadia G, Frank J, Sauerbier M, Neary P, Gallagher AG, Traynor O and Carroll S., "Factors influencing microsurgical skill acquisition during a dedicated training course," *Microsurgery*, 32(8): 649–56, 2012.
- [2] Hong JW, Kim YS, Lee WJ, Hong HJ, Roh TS and Song SY., "Evaluation of the efficacy of microsurgical practice through time factor added protocol: microsurgical training using nonvital material," *The Journal of Craniofacial Surgery*, 21(3): 876–881, 2010.
- [3] Matsumura N, Horie Y, Shibata T, Kubo M, Hayashi N and Endo S., "Basic training model for supermicrosurgery: a novel practice card model," *Journal of Reconstructive Microsurgery*, 27(6): 377–382, 2011.
- [4] Gallagher AG, Richie K, McClure N, and McGuigan J., "Objective psychomotor skills assessment of experienced, junior, and novice laparoscopists with virtual reality," *World Journal of Surgery*, 25(11): 1478–1483, 2001.
- [5] McClusky DA, Ritte EM, Lederman AB, Gallagher AG, Smith CD, Park AE, Scott D, and McClusky DA., "Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator," *American Surgeon*, 71(1): 13–21, 2005.
- [6] Herzele IV, O'Donoghue KGL, Aggarwal R, Vermassen F, Darzi A, and Cheshire NJW., "Visuospatial and psychomotor aptitude predicts endovascular performance of inexperienced individuals on a virtual reality simulator," *Journal of Vascular Surgery*, 51(4): 1035–1042, 2010.
- [7] Murdoch JR, Bainbridge LC, Fisher SG, and Webster MHC., "Can a simple test of visual-motor skill predict the performance of microsurgeons?," *Journal of the Royal College of Surgeons of Edinburgh*, 39(3): 150–152, 1994.
- [8] Kent RN., and Foster SL., "Direct observational procedures: Methodological issues in naturalistic settings," *Handbook of behavioral assessment*. New York: Wiley, pp. 279–328, 1977.
- [9] Carraccio C, and Englander R., "The objective structured clinical examination: a step in the direction of competency-based evaluation," *Archives of Pediatrics & Adolescent Medicine*, 154(7): 736–741, 2000.
- [10] Sidhu RS, Grober ED, Musselman LJ, and Reznick RK., "Assessing competency in surgery: where to begin?," *Surgery*, 135(1): 6–20, 2004.
- [11] Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, and Brown M., "Objective structured assessment of technical skill (OSATS) for surgical residents," *British Journal of Surgery*, 84(2): 273–278, 1997.
- [12] Faulkner H, Regehr G, Martin J, & Reznick R., "Validation of an objective structured assessment of technical skill for surgical residents," *Academic Medicine*, 71(12): 1363–1365, 1996.
- [13] Goff BA, Lentz GM, Lee D, Fenner D, Morris J, and Mandel LS., "Development of a bench station objective structured assessment of technical skills," *Obstetrics and Gynecology*, 98(3): 412–416, 2001.
- [14] Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, Murakami S, Saeki S, Mukaida H, and Takiyama W., "Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room," *Surgery Today*, 43(3): 271–275, 2013.
- [15] Doyle JD, Webber EM, and Sidhu RS., "A universal global rating scale for the evaluation of technical skills in the operating room," *American Journal of Surgery*, 193(5): 551–555, 2007.
- [16] Chan W, Niranjan N, and Ramakrishnan V., "Structured assessment of microsurgery skills in the clinical setting," *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 63(8): 1329–1334, 2010.
- [17] Reiley CE, Lin HC, Yuh DD, and Hager GD., "Review of methods for objective surgical skill evaluation," *Surgical Endoscopy*, 25: 356–366, 2011.
- [18] Ahmmad SNZ, Su ELM, Fai YC, and Khairi F., "Assessment methods for surgical skill," *World Academy of Science, Engineering and Technology*, 58: 752–758, 2011.
- [19] Ahmed K, Miskovic D, Darzi A, Athanasiou T, and Hanna G B., "Observational tools for assessment of procedural skills: a systematic review," *The American Journal of Surgery*, 202(4): 469–480, 2011.
- [20] Khalifa YM, Bogorad D, Gibson V, Peifer J, and Nussbaum J., "Virtual reality in ophthalmology training," *Survey of Ophthalmology*, 51(3): 259–273, 2006.
- [21] Kazemi H, Rappel JK, Poston T, Hai Lim B, Burdet E, and Leong Teo C., "Assessing suturing techniques using a virtual reality surgical simulator," *Microsurgery*, 30(6): 479–486, 2010.

- [22] Lam CK, Sundaraj K, Sulaiman MN, & Qamaruddin FA., "Virtual phacoemulsification surgical simulation using visual guidance and performance parameters as a feasible proficiency assessment tool," *BMC Ophthalmology*, 16(1): 1-9, 2016.
- [23] Ahmmad SNZ, Su ELM, Yeong CF, and. Narayanan ALT., "Experimental study of surgeon's psychomotor skill using sensor-based measurement," *Procedia Computer Science*, 42:130-137, 2014.
- [24] S Ahmmad SNZ, Su ELM, Yeong CF., "Force variability as an objective measure of surgical skill," *Jurnal Teknologi*, 74(6): 125-128, 2015.
- [25] Lam CK, Sundaraj K, and Sulaiman MN., "Virtual reality simulator for phacoemulsification cataract surgery education and training," *Procedia Computer Science*, 18: 742-748, 2013.
- [26] Grober ED, Hamstra SJ, Wanzel K R, Reznick RK, and Matsumoto ED., "Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity," *Microsurgery*, 23(4): 317-322, 2003.
- [27] Saleh GM, Gauba V, Sim D, Lindfield D, Borhani M, and Ghoussayni S., "Motion analysis as a tool for the evaluation of oculoplastic surgical skill: evaluation of oculoplastic surgical skill," *Archives of Ophthalmology*, 126(2): 213-216, 2008.
- [28] Grober ED, Roberts M, Shin E J, Mahdi M, and Bacal V., "Intraoperative assessment of technical skills on live patients using economy of hand motion: establishing learning curves of surgical competence," *The American Journal of Surgery*, 199(1): 81-85, 2010.
- [29] Benaoumeur I, Zoubir A, and Reda HEA., "Remote Control of Mobile Robot using the Virtual Reality," *International Journal of Electrical and Computer Engineering*, 5(5): 1062-1074, 2015.
- [30] Xi-Lin K, Qing-Sheng G, Yue-Peng Z, Yan Z., "Construction and Application of Virtual Reality Geographic Information System," *Indonesian Journal of Electrical Engineering and Computer Science*, 12(6): 4353-4360, 2014.
- [31] Zhu Y., "QuickTime Virtual Reality Technology Applies to Practical Teaching Recording System," *Indonesian Journal of Electrical Engineering and Computer Science*, 11(11):6315-6320, 2013.
- [32] Kazemi H, Rappel JK, Poston T, Hai Lim B, Burdet E, and Leong Teo C., "Assessing suturing techniques using a virtual reality surgical simulator," *Microsurgery*, 30(6): 479-486, 2010.
- [33] Beers RJ., Sittig AC, and van der Gon Denier JJ., "How humans combine simultaneous proprioceptive and visual position information," *Experimental Brain Research*, 111(2): 253-261, 1996.
- [34] Ahmmad SNZ, Su ELM, Yeong CF, Sood S, and Gandhi A., "Objective Measurement for Surgical Skill," *Jurnal Teknologi*, 78(7-5): 145-152, 2016.
- [35] Wong JD, Wilson ET, Kistemaker DA, and Gribble PL., "Bimanual proprioception: are two hands better than one?," *Journal of Neurophysiology*, 111(6): 1362-1368, 2014.
- [36] Hofstad EF, Våpenstad C, Chmarra MK, Langø T, Kuhry E, and Mårvik R., "A study of psychomotor skills in minimally invasive surgery: what differentiates expert and nonexpert performance," *Surgical Endoscopy*, 27(3): 854-863, 2013.
- [37] Su ELM, Ganesh E, Yeong CF., and Burdet E., "Accurate micromanipulation induced by performing in unstable dynamics," *19th IEEE International Symposium on Robot and Human Interactive Communication*, 762-766, 2010.
- [38] Risucci DA.NVisual spatial perception and surgical competence," *American Journal of Surgery*, 184: 291-295, 2002.
- [39] Sturman MM, Vaillancourt DE, and Corcos DM., "Effects of aging on the regularity of physiological tremor," *Journal of Neurophysiology*, 93: 3064-3074, 2005.