World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:5, No:10, 2011

Assessment Methods for Surgical Skill

Siti Nor Zawani Ahmmad, Eileen Su Lee Ming, Yeong Che Fai and Fauzan Khairi bin Che Harun

Abstract**—**The increasingly sophisticated technologies have now been able to provide assistance for surgeons to improve surgical performance through various training programs. Equally important to learning skills is the assessment method as it determines the learning and technical proficiency of a trainee. A consistent and rigorous assessment system will ensure that trainees acquire the specific level of competency prior to certification. This paper reviews the methods currently in use for assessment of surgical skill and some modern techniques using computer-based measurements and virtual reality systems for more quantitative measurements

Keywords**—**assessment, surgical skill, checklist, global rating, virtual reality

I. Introduction

THE complexity of the microsurgery techniques requires high dexterity maneuvering, and thus a lot of training from surgeons [1]. The skill level will change and vary after they go through training, and gain experience during surgical practice. A well-thought assessment system can help to identify the level of technical skill and competence among the surgeons as well as the trainees as they participate in the training program. Assessment forms the basis of education and training process. The UK Postgraduate Medical Education and Training Board defines assessment as the process to measure progress and achievement of an individual against defined standards and criteria, which often includes an experimental measurement [2]. In the educational context, the purpose of assessment is to measure the improvement of skills and knowledge from time to time, to rank the people for the selection and also can motivate them for better performance [3]. The need of early assessment is important in medical profession in order to identify trainees suitable for the profession and to prevent any potential bias happening in the operating room. Based on the outcome of assessment, then a more personalized training can be given and tailored to the actual skill level of the individual in order to achieve technical proficiency. This paper reviews some assessment methods currently in use to grade surgical skill and the potential of adding computer-based measurements to complement conventional methods.

II. Method of Literature Search

All the source was obtained from the years 1999 through 2011 with the combination keywords such as “microsurgery technique assessment”, “technical skill”, “surgery”, Only peer-reviewed journal articles and book chapters are selected. Additional source included articles cited in the references list of others articles.

All authors are with the Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81300 UTM Johor Bahru, Johor, Malaysia. (phone: +607-5535237; fax: +607-5566272; e-mail: eileensu@fke.utm.my).

III. Surgical Skill Assessment

A. Traditional Assessment

In the traditional method, assessment in the operating room is done based on direct observation by expert surgeons. This assessment is extremely subjective, not criteria-based, and can be unreliable as it is affected by inter-observer variability and recall bias [4]. Besides that, the evaluators need to spend a large amount of time in order to observe the performance of technical skill in the operating room.

Logbook is another method commonly used in the United Kingdom. It recorded the trainee experience but not the performance within an operating theatre. The lack of information about their proficiency cause this method to be subjective and having poor validity. Surgical assessment needs to be made more objective to provide quality and more accurate feedback during the process of surgical training.

B. The Global Rating index for Technical Skills (GRITS)

The Global Rating index for Technical Skills (GRITS) was designed based on the Global rating Scale (GRS). This assessment tool can be used in open and laparoscopic surgery. The observation by expert with the aid of fixed criteria makes the assessment more reliable, valid and feasible when the evaluation tool is implemented in the operating room during general surgery cases. Doyle et al selected 9 items to be general markers of technical skill, and their measurements are by a rating scale from 1 to 5 [4]. The system was tested with seven general surgery residents of University British Columbia from various postgraduate years (PGY) 1-6. They were evaluated by 13 teaching faculties at 4 different hospitals. The trainees’ performance was for a wide variety of cases. All residents are being evaluated compared with the gold standard. The result shows that GRITS is able to distinguish between resident of different PGY levels [4]. GRITS tool is easy to use and the cost on the examiner time is minimal. It also provides helpful feedback for residents who want to enhance and develop their technical skill continuously. 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 biased based on the background level of trainee. For example, the evaluator tends to avoid giving the lower end of the scale. This assessment tool can be enhanced if the observer was someone who is not the trainee’s preceptor.

TABLE I
GLOBAL RATING INDEX FOR TECHNICAL SKILLS (GRITS) [4]

Respect for tissue				
1	2	3	4	5
Frequent unnecessary force on tissue or caused damaged by inappropriate use of instrument		Careful handling tissue but occasionally caused inadvertent damage		Consistently handled tissue appropriately with minimal damages of tissue
Time and motion				
1	2	3	4	5
Many unnecessary moves		Efficient time/motion but unnecessary moves		Clear economy of movement. Maximum efficiency
Instrument handling and knowledge				
1	2	3	4	5
Tentative/awkward moves or inappropriate use		Competent use of instrument		Fluid moves wih instrument. No awkwardness
Flow of operation				
1	2	3	4	5
Frequently stopped, seemed unsure of next move		Some forward planning, reasonable progression		Obviously planned course, effortless flow
Knowledge of specific procedure				
1	2	3	4	5
Deficient knowledge. Required specific instruction at most steps.		Knew all important steps of operation		Demonstrated familiarity with ll steps of operation
Use of assistants (if applicable)				
1	2	3	4	5
Consistently placed assistants poorly or failed to use		Appropriate use of assistants most of the time		Strategically used assistants to best advantages to all time.
Communication skills				
1	2	3	4	5
Frequent problems working with team or fail to communicate		Appropriate communicatio n with team most of the time.		Co-ordinates surgical team in a superior manner
Depth perception (Laparoscopic only)				
1	2	3	4	5
Constantly overshoots, swings wide, slow correction		Some overshooting but quick to correct		Accurately directs instruments in correct plane
Bimanual Dexterity				
1	2	3	4	5
Use only one hand, poor coordination between hands		Use both hands but does not optimize their interaction		Expertly uses both hands to provide optimal exposure

C. Sructured Assessment of Microsurgery Skills in the clinical Setting (SAMS)

WoanYi et al. proposed the SAMS method for clinical setting which combine three types of assessment; modified Global Rating Score (GRS), error list, and summative rating [2]. SAMS method was used to assess the performance of trainees and consultant handling fifteen clinical cases of microvascular anastomoses. All the performances were recorded on video using a digital microscope system. These recordings were viewed by three consultants independently in blinded fashion and marks were given based on the subject's performance by referring to the twelve items of modified GRS. Those items were grouped by four main areas of microsurgery skill as shown in Table 2.

The error list in the SAMS method is used to facilitate the evaluator to identify errors made by subject. Summative rating was then applied to give the summative feedback on the overall performance and to identify the level of skill attained. The result shows that the consultant had the highest scores in all the technical components of the modified GRS, in the summative scores and no errors as rated by all three assessors while the trainees had a lower score in most parameter.By using video analysis, it provides the qualitative method of assessment and performance can be quantified through the scale points. However, some microsurgery skill parameter is difficult to assess through video such as steadiness. In order to overcome this problem, the researcher suggested the use of computerized system such as hand motion analysis and virtual reality. Using sensors, these motion parameters could be more accurately measured.

D. Patient Robot

Researchers at Waseda University have proposed the development of the patient robot as an advanced evaluation tool for medical procedure as it will provide more detailed information of the task. The WKS-2RII have been designed, which consisted a personel computer, a web cam, and a dummy skin model with embedded sensors [5]. The dummy skin was made from polyurethane rubber which has almost the same sensation as human skin.

TABLE II
SAMS PARAMETERS TO EVALUATE ANASTOMOSIS PERFORMANCE [2]

Parameter	Explanation
Dexterity -steadiness -instrument handling -tissue handling	Dexterous handling of tissue is important to minimize tissue damage and reduce the risk of thrombosis. It is a basic prerequisite for microsurgery.
Visuo-spatial ability -dissection -suture placement -knot technique	Familiarization on performing surgery under the microscope is very important. Well-prepared vessel ends improve visualization of the vessel walls and ensure accurate suture placement. The placement and spacing of sutures require visuo-spatial awareness to avoid catching the back-wall and suture entanglement. Knot technique and tightening under the microscope is done under vision rather than by feel.
Operative flow -Step -Motion -Speed	This relates to the whole procedure of completing an anastomosis efficiently. Knowledge and control of the steps are important factors to determine the operative flow. The control of each movement (motion) contributes to an efficient operative flow of the procedure.
Judgement -irrigation -patency test -Bleeding control	Judgment is the ability to recognize, prevent and manage complication. It requires irrigation, patency test and bleeding control. Irrigation can help distending the lumen and prevent tissue desiccation, but overuse can increase surface tension causing suture adherence. A patency test needs to be done delicately to avoid injury to the intima. Adequately placed sutures through judgment can avoid anastomosis leak and so control bleeding.

The embedded sensors are used to collect information about the movement of the skin when a trainee performs the suture/ligature task. The performance of the trainee is recorded using the web cam. They also proposed the suture image-processing algorithm in order to evaluate the physical properties of the suture after finishing the task. For the assessment, six parameters were evaluated as shown in Table 3, by considering the objective structured clinical examination (OSCE) examination's checklist.

TABLE III
OSCE PARAMETERS TO EVALUATE SUTURING [5]

Parameter	Explanation
Time (T)	Conventionally, the time is measured from start, up to the completion of the task. Less time is required for the surgeon to obtain a better performance (with precise movements achieved).
Force in Tissue (FoT)	Insertion of needles can damage the skin if the trainee applies shear force. It will be less damaging if insertion is done along the trajectory of needle curve. In the WKS-2RII, the *Force in Tissue* parameter is determined by measuring the sideway movements of dummy skin.
Judging Tension (JuT)	*Judging Tension* parameter is determined by measuring the deformation of dummy skin after ligature. It might be painful for the patient if the ligature is done too tightly and also can cause infection around the wound area if not properly done.
Equidustance (EqD)	The suture width is measured between the needle insertion point and the wound edge. This parameter is determined by identifying the location of each suture and computing its width.
Distance between Suture (DbS)	The suture distance is measured from one suture to another applied along the wounded area. It is determined by identifying the center of mass of each suture and computing the distance between sutures.
Wound Dehiscence (WoD)	This parameter is quantified by measuring the open wound area after suturing. Smaller open area of the wound indicates better performance.

By using discriminant analysis method, the weighting coefficient of each parameter can be determined and implemented through the experiment. The experimental data was collected from surgeons, medical student and unskilled person.

Result shows that the proposed parameter can differentiate between different levels of expertise. This system can be used as an educational tool where an unskilled person can be trained and assessed. Possible risks on patients can be reduced due to the use of non-living model.

E. Hand Motion Analysis

Fig. 1Electromagnetic trackers placed on backs of mi crosurgeons' hands allowed positional data to be recorded and analyzed by computer software while subject performed a standardized microsurgical task [9]

Fig. 2 Hand-motion analysis was recorded using Imperial College surgical assessment device (ICSAD) [9]

The Imperial College Surgical Assessment Device is a dexterity-based motion analysis device developed by the Department of Surgical Oncology and Technology by the Surgical Computing and Imaging Research Group [6]. The technology using electromagnetic tracking system (Isotrak II; Polhemus Inc, Colchester, Vt) is connected to a portable computer through a standard RS-232 (serial) port, independent motion acquisition software, and custom-made analysis software to track the surgeon's hand motion when performing a surgical task [6]. Two sensors are attached on the dorsal surface of the surgeon's hand as shown in Fig. 1. The positional data from trackers are recorded and analyzed in order to determine the number of hand movements, hand travel distance, direction, acceleration changes as well as the time taken to complete the task [7],[8]. This device has been shown to be a valid quantitative measure of dexterity in laparoscopic and open surgical simulation. Grober et al used this hand motion analysis as an objective measures of microsurgical performance [9]. The overall system is shown in Fig. 2. They asked the subjects to complete the several baseline microsurgical was assessed by expert using global

TABLE IV
SUMMARY OF MOMS TASKS [10]

Station	Explanation	Method
Operative equipment and instrument	This station tested the knowledge of the subject. It involved instrument and operating theatre equipment where the subject had to select the most appropriate tool based on the task given. They answered questions related to the tool.	Examination
Knot tying	The subject tied suture knot based on the task. Two parameters were analysed: number of hand movement made and time taken to complete task.	Motion analysis
Suturing	Two standardized exercises of suturing were involved: -five simple interrupted suture with 4 single instrument throws; -five interrupted vertical mattress sutures with 4 single instrument throws. Time taken to complete task and number of movements made were measured.	Motion analysis
Closure of enterotomy	This involved the closure of a 2-cm small bowel enterotomy using 2 stay sutures and interrupted seromuscular sutures as advocated on the Basic Surgical Skills Course. Performance was recorded using camera and reviewed by three expert surgeons.	Objective Structured Assessment of Technical Skills (OSATS) IV.
Excision of sebaceous cyst	This required the excision of a synthetic sebaceous cyst which consisted of a multilayered foam pad surrounding a polymer capsule containing a yellow oily liquid. It has the potential to burst if incised. Recording of performance was reviewed by expert surgeon.	OSATS
Laparoscopic task	Two tasks were carried out: - Single handed task (Grasping an object and placing at wire frame cage) - Traversal (2-handed task)	MIST-VR simulator

rating scale which consists six parameters and each will grade the performance on a five point scale. After the training, hand motion analysis scores improved significantly [9]. The study shows that hand motion analysis is an objective, valid, and sensitive method for assessing technical skill where potential examiner bias can be eliminated. This system can be adapted to a variety of procedure and the presence of expert examiner is not necessarily required. However, the use of electromagnetic tracker on dorsal hand may interfere surgeons' movement. One device may only access one candidate; therefore to complete an assessment on a class of participants will involve more hardware costs.

A. Multiple Objective Measures of Surgery (MOMS)

Bann conducted as study to validate the multiple objective measures of surgery (MOMS) method. MOMS was designed to assess the basic level of surgical skill among the trainees objectively using combination of six-bench top station [10]. Various assessment criteria were included, such as motion analysis, observation with criteria and inbuilt simulation metric. Two different levels of surgical trainees were tested; basic surgical trainees (BSTs) and higher surgical trainees (HSTs) who had at least two years' experience at this level. The summary of the tasks is shown in Table 4. Results of MOMS show that the assessment is reliable and valid, and can be used for the assessment of basic technical skill.

B. Virtual Reality

Virtual reality (VR) is defined as a collection of technologies that allow people to interact efficiently with three dimensional (3D) computerized database techniques in real time simulation using the natural sense and skills. The process involves graphic computer which provide a 3D graphic images. Virtual reality technology is slowly gaining acceptance by a surgeons for training purposes. Several VR simulators have been developed to train and assess surgical skills. Ziegler et al developed an arthroscopy training simulator by applying computer graphics system and virtual reality technique [11]. The 3D interaction of the system simulates a real arthroscopy case. A tracking system is used in the arthroscopy simulator in order to get the position and orientation of the tibia and the instrument. It consists of three tracker units, one for the transmitter which is located underneath the femur, one on the arthroscope, and another one is in tibia. Fig. **3** Hardware configuration of the VR arthroscopy simulator

shows the hardware configuration of the VR arthroscopy simulator. The prototype of this simulator have been successfully presented, tested and assessed during arthroscopy training course at the "Frankfurt Sport Medicine Weekend". The Minimally Invasive Surgical Trainer (MIST) developed in the UK was the first VR simulator used as an assessment tool. In the study by Chaudhry et al, MIST VR, a PC-based laparoscopic cholecystomy simulator provides objective assessment of psychomotor skills based on errors made and time taken for six different tasks [12]. It gives an overall score for performance which is shown to be able to distinguish between experienced and non-experienced surgeon.

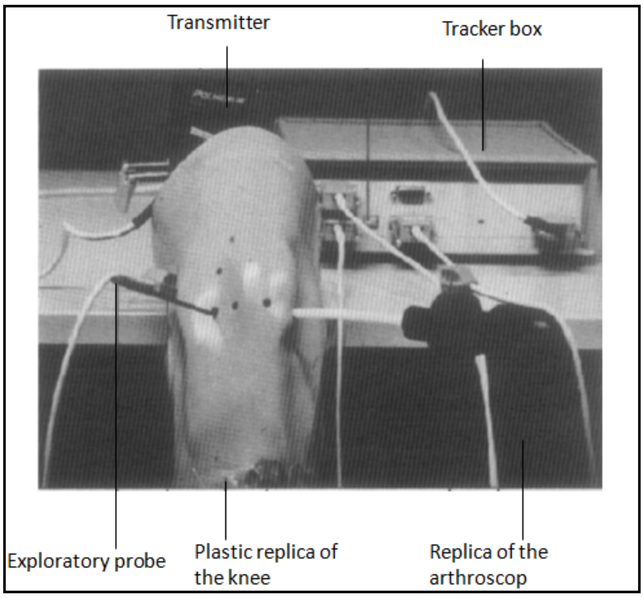

Fig. 3 Hardware configuration of the VR arthroscopy simulator

Simbionix Uromentor endoscopic simulator is a VR device used for practising the basic endoscopic skill for novice endoscopists, shown as Fig.4. By using uromentor program, the number of parameter such as total procedure time, time to insert guidewire, incidence of mucosa trauma from the instruments, number of perforations and fragmentation time can be assessed. This assessment was evaluated by two experienced surgeon using global rating scale. The study led by Wilhelm showed that students who trained on the VR simulator gained improvement in endoscopic skills. This training allows a student to familiarize with instrumentation and also can enhance educational experience among the students [13].

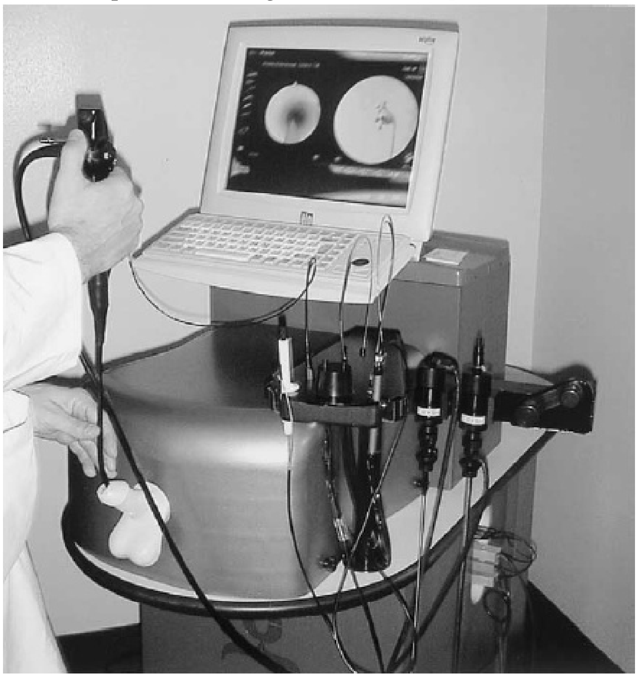

Fig. 4 Simbionix Uromentor Virtual Reality Simulator

The GI Mentor II (Simbionix Ltd, Cleveland, OH, USA) consisted of mannequin with body orifices for upper/lower endoscopy, a Pentax ECS-3840F endoscope, guide wire devices and foot pedals. The metric of performance included time to complete the task, efficiency and several observational parameters such as percentages of pathology seen. All the parameters were calculated and registered by the computer system. Significant differences were seen for all metrics between the novice and experienced groups. These results showed that the simulator can differentiate between subject with different level of endoscopic skill (experienced and novices). Experienced subjects were faster, having higher percentage of the mucosa, had fewer 'red-out' and caused less discomfort for virtual patient compared to novices [14].

The da Vinci Trainer is a virtual reality simulator based on Mimic Technologies' Mantis Duo platform, a 2-handed haptic system that fits on a tabletop and shares a common lineage with the University of Washington transurethral resection of the prostate simulator [15]. Several simulation modules are available in this system such as camera movement, targeting, robotic arm movement, object manipulation and suturing. It can be used to train and assess surgeons in the field of laparoscopic surgery. Kenney et al asked novice and experienced surgeon to complete two modules namely the EndoWrist module ("Pick and Place" and "Peg Board") and needle driving module ("Dots and Number" and "Suture Sponge") as shown in Fig.5. Then, performance of the subjects was measured using a computerized built-in scoring algorithm. The percentages of the total score was derived from the combination of the selected variable such as total task time, maximal force and strain, total instrument motion, number of instrument collision, time instrument were out in view, number of target attempted successfully and not attempted. The result shows that, this simulator was able to differentiate between novice and expert robotics surgeon [15].

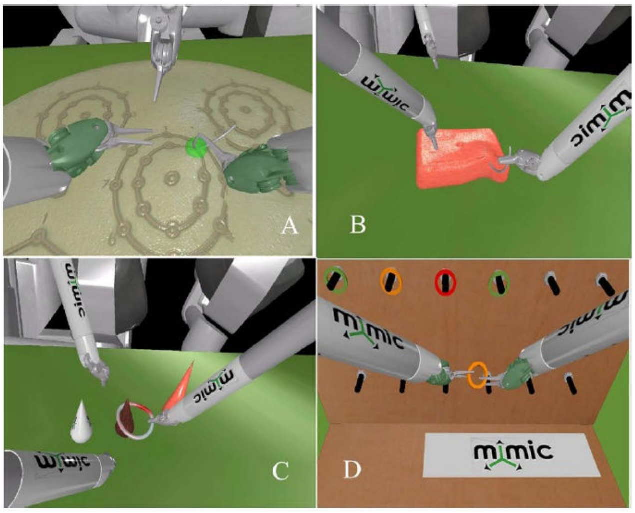

Fig. 5 Screen images from 4 tested modules: (A) Dots and Numbers, (B) Suture Sponge, (C) Pick and Place, and (D) Peg Board [15]

V. CONCLUSION

The variety of assessment methods, be it visual observation, computerized measurements or their combination, show some level of validity in differentiating experienced surgeon from novices. These assessments provide helpful feedback for residents going through technical skill development. Some important points to be considered for a successful assessment include validity, reliability and feasibility. Assessment is a crucial step in the learning process; therefore an accurate and comprehensive assessment of the technical skills will facilitate skill acquisition, identify the requisite attributes in trainees and test essential medical knowledge prior to certification. Much work is still being done to improve surgical skill assessment, by using sensors and mathematical algorithms besides expert rating, to make it fair, robust and comprehensive for various surgical applications.

ACKNOWLEDGMENT

This work is supported by Universiti Teknologi Malaysia research grant 03J68 and 4P006.

REFERENCES

[1] E. Erel, B. Aiyenibe, and P. E. M. Butler, "Microsurgery simulatorin virtual reality:Review," *Microsurgery,* vol. 23, pp. 147-152, 2003.

[2] N. N. WoanYi Chan, Venkat Ramakrishnan, "Structured assessment of microsurgery skills in the clinical setting," *Journal of Plastic, Reconstructive & Aesthetic Surgery,* vol. 63, pp. 1329-1334, 2010.

[3] R. M. Epstein and E. M. Hundert, "Defining and assessing professional competence," *Journal of American Medical Association,* vol. 287, pp. 226-235, 2002.

[4] J. D. Doyle, Eric M. Webber, and R. S. Sidhu, "A universal global rating scale for the evaluation of technical skills in the operating room," *The American Journal of Surgery* vol. 193, pp. 551-555, 2007.

[5] J. Solis, N. Oshima, H. Ishii, N. Matsuka, K. Hatake, and A. Takanishi, "Development of a sensor system towards the acquisition of quamtitative information of the training progress of surgical skills," in *Proceedings of the 2nd Biennial IEEE/RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics* Scottsdale, AZ, USA, 2008, pp. 19-22.

[6] A. Dosis, R. Aggarwal, F. Bello, K. Moorthy, Y. Munz, D. Gillies, and M. Ara Darzi, "Synchronized Video and Motion Analysis for the Assessment of Procedures in the Operating Theater," *Arch Surg.,* vol. 140, pp. 293-299, 2005.

[7] P. U. Kalu, J. Atkins, D. Baker, C. J. Green, and P. E. M. Butler, "How do we assess microsurgery skill? ;:," *Microsurgery,* vol. 25, pp. 25-29, 2005.

[8] 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," *American College of Surgeons,* vol. 193, pp. 479-485, 2001.

[9] E. D. Grober, S. J. Hamstra, R. K. Reznick, E. D. Matsumoto, R. s. Sidhu, and K. A. Jarvi, "Validation of novel and objective measures of microsurgical skill: Hand-motion analysis and stereoscopic visual acuity," *Microsurgery,* vol. 23, pp. 317-322, 2003.

[10] S. Bann, I. M. Davis, K. Moorthy, Y. Munz, J. Hernandez, M.Khan, V. Datta, and A. Darzi, "The Reliability of multiple objective measures of surgery and the role of human performance," *The American Journal of Surgery,* vol. 189, pp. 747-752, 2005.

[11] R. Ziegler, G. Fischer, W. Muller, and M. Gobel, "Virtual reality arthroscopy training simulator," *Comput. Biol. Med,* vol. 25, pp. 193-203, 1995.

[12] A. Chaudhry, C. Sutton, J. Wood, R. Stone, and R. McCloy, "Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface," *Ann R Coll Surg Engl,* vol. 81, pp. 281-286, 1999.

[13] D. M. Wilhelm, K. Ogan, C. G. Roehrborn, J. A. Cadeddu, and M. S. Pearle, " Assessment of Basic Endoscopic Performance Using a

Virtual Reality Simulator," *Journal of the American College of Surgeons,* vol. 195, pp. 675-681., 2002.

[14] T. P. Grantcharov, L. Carstensen, and S. Schulze, "Objective Assessment of Gastrointestinal Endoscopy Skills Using a Virtual Reality Simulator," *Journal of the Society of Laparoendoscopic Surgeons,* vol. 9, pp. 130-133, 2005.

[15] P. A. Kenney, M. F. Wszolek, J. J. Gould, J. A. Libertino, and A. Moinzadeh, "Face, Content, and Construct Validity of dV-Trainer, a Novel Virtual Reality Simulator for Robotic Surgery," *Urology,* vol. 73, pp. 1288-1292, 2009.