Published in final edited form as:

Laryngoscope. 2014 October ; 124(10): 2332–2337. doi:10.1002/lary.24724.

Video-Based Method of Quantifying Performance and Instrument Motion During Simulated Phonosurgery

Ellen Conroy, Ketan Surender, MSc, Zhixian Geng, PhD, Ting Chen, MD, Seth Dailey, MD, and Jack Jiang, MD, PhD

Department of Surgery (E.C., K.S., Z.G., S.D., J.J.), Division of Otolaryngology–Head and Neck Surgery, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, U.S.A.; Department of Otolaryngology–Head and Neck Surgery (T.C.), Fujian Provincial College of Fujian Medical University, Fuzhou, China

Abstract

Objectives/Hypothesis—To investigate the use of the Video-Based Phonomicrosurgery Instrument Tracking System to collect instrument position data during simulated phonomicrosurgery and calculate motion metrics using these data. We used this system to determine if novice subject motion metrics improved over 1 week of training.

Study Design—Prospective cohort study.

Methods—Ten subjects performed simulated surgical tasks once per day for 5 days. Instrument position data were collected and used to compute motion metrics (path length, depth perception, and motion smoothness). Data were analyzed to determine if motion metrics improved with practice time. Task outcome was also determined each day, and relationships between task outcome and motion metrics were used to evaluate the validity of motion metrics as indicators of surgical performance.

Results—Significant decreases over time were observed for path length ($P < .001$), depth perception ($P < .001$), and task outcome ($P < .001$). No significant change was observed for motion smoothness. Significant relationships were observed between task outcome and path length ($P < .001$), depth perception ($P < .001$), and motion smoothness ($P < .001$).

Conclusions—Our system can estimate instrument trajectory and provide quantitative descriptions of surgical performance. It may be useful for evaluating phonomicrosurgery performance. Path length and depth perception may be particularly useful indicators.

Keywords

Phonomicrosurgery; surgical assessment; surgical training; simulated model; Video-Based Phonomicrosurgery Instrument Tracking System

Send correspondence to Jack J. Jiang, MD, 1300 University Avenue, 2745 Medical Sciences Center, Madison, WI 53706. jjjiang@wisc.edu.

Level of Evidence: NA

The authors have no funding, financial relationships, or conflicts of interest to disclose.

INTRODUCTION

Ensuring surgeon competency is vital for providing safe, effective healthcare. Traditional methods of surgical assessment include logbooks, direct observation, simulated models with scoring criteria, review of video-recorded procedures, psychomotor testing, and motion analysis.¹⁻³ Few of these approaches offer objective and quantitative performance assessments. Meaningful evaluation of resident progress through surgical training and surgeon credentialing requires objective assessment that eliminates common grading bias.^{2,4} Structured human grading systems attempt to achieve standardization through rated checklists on inanimate benchtop models.^{1,2,5} Still, this method requires expert personnel, remains based on observation, is subject to a shortage of monitored training centers, needs well-structured task stations, and is time consuming.^{1,4-7}

Traditional methods, based on observation by senior faculty, rarely include descriptions of instrument motion.¹ Hand and instrument motion analysis has been successfully used in several surgical fields to provide an analysis of surgical dexterity, a crucial component of surgeon competency.^{1,7-9} Recently, hand motion analysis has been used to differentiate between oculoplastic surgeons of differing experience levels, demonstrate skill improvement after a microsurgical skills course, and assess technical skill during live vasectomy and vasectomy reversal.^{7,10,11} Motion scores improved over time for novice surgeons, but remained stable for experienced surgeons.¹¹ Instrument motion analysis has also been used to differentiate between expert and novice laparoscopic surgeons.¹² Furthermore, motion analysis scores correlate with global rating and checklist assessment scores, such as the Objective Structured Assessment of Technical Skills, a widely accepted tool for assessing surgical skill.¹³ Similarly, path length, derived from video motion analysis during simulated endovascular surgery, correlates with qualitative performance scores.¹⁴

Several studies cite the challenge of determining which surgical task may best differentiate between surgical ability levels. Certain tasks may more heavily evaluate a specific component of dexterity; surgical procedures of greater length and complexity were not evaluated.^{10,11} Expense and availability of motion analysis have also been noted as obstacles to its widespread use.^{11,15} As quantitative motion analysis data do not reveal the quality of surgical outcome, the use of global rating scales is recommended in conjunction with motion analysis.¹⁵ Surgical ability theoretically lies within motion structure, which is the result of elements that affect technical abilities including dexterity, instrument control, and environmental factors.⁵ Motion metrics computed from position data that are considered to be important in microsurgery and endoscopic surgery are path length, number of movements, motion smoothness, and depth perception.^{5,7,16,17}

Such methods of performance evaluation may be useful for phonomicrosurgery, a surgical subspecialty with numerous intraoperative challenges, including the use of instruments approximately 30-cm long and a small operative area in front of and below the surgeon's vision, requiring surgeons to extend their arms forward during operations. These factors increase fatigue, discomfort, and gravity-induced tremor.¹⁸⁻²⁰ The relatively recent emergence of this subspecialty and its distinctive physical requirements warrant a method capable of accurately evaluating surgical performance. To our knowledge, video-based

motion analysis has not been purposed for use in evaluating this subspecialty. Furthermore, very few studies have focused on education and training in phonosurgery; using a laryngeal dissection station to establish clinical competency while maintaining low costs has been proposed.^{21,22} Utilizing motion analysis in conjunction with this station would allow for tracking progress and could give valuable feedback to residents during training without endangering patients. This will improve surgeon evaluation and training, clinical care, and research. Progress in this subspecialty will reduce the impact of voice disorders, decrease iatrogenic morbidity, and increase positive treatment outcomes, leading to improved patient quality of life.

We have created a new method of measuring instrument movement in a simulated surgical environment, termed the Video-Based Phonomicrosurgery Instrument Tracking System (V-PITS). V-PITS consists of a data acquisition and algorithm component. The former consists of hardware and software needed to capture time-synchronized video from two different viewpoints. The latter consists of software that processes synchronized videos to compute three-dimensional instrument motion information. This system does not require attachment of wired sensors that may interfere with hand or instrument motion, or placement of instruments into a mechanical rig to measure movement. The use of video analysis eliminates concern about interference due to instrument or equipment material, which can be problematic for electromagnetic- and infrared-based systems. Furthermore, the ultimate goal in surgeon evaluation may be to track surgical instruments using cameras only, as this introduces minimal additional hardware into the operating room. Although V-PITS is not currently ready for integration into the operating room, it is a start in moving toward this goal. Finally, V-PITS is tailored toward phonomicrosurgery, a unique and challenging subspecialty.

Instrument motion was measured during a program of daily simulated surgical tasks over 5 days to determine the effectiveness of V-PITS to collect motion data as an objective and quantifiable evaluation of phonomicrosurgical performance. Global rating scales for quality of surgical outcome are generally better suited to longer, more complex surgical procedures and were impractical for the task we used. Instead, we used task outcome scores based on deviation of subject cut lines from the target line. To evaluate V-PITS, we investigated relationships between task outcome and quantitative motion metrics, and change in surgical performance over time. We hypothesized that increased practice time would result in decreased task outcome scores and motion metrics, including path length, depth perception, and motion smoothness, and that task outcome would have a significant relationship with each motion metric.

MATERIALS AND METHODS

Data Collection

V-PITS—V-PITS measures three-dimensional instrument motion using a stereo camera rig consisting of two Basler acA640-100gc machine vision cameras (Basler Vision Technologies, Ahrensburg, Germany) equipped with Fujinon 25-mm c-mount lenses (Fujinon Corp., Saitama, Japan). Two synchronized digital video-recordings of instrument tip movement during simulated surgical tasks are captured. An Agilent 33220A (Agilent

Technologies, Santa Clara, CA) controls frame capture rate and provides synchronization between cameras. A custom software application written in C++ that relies on Basler Pylon SDK for Windows (Basler Vision Technologies) and Open Source Computer Vision Library (Willow Garage, Menlo Park, CA) is used to acquire video. A frame rate of 30 frames per second was used.

To compute motion information, synchronized videos are processed offline using custom software written in MATLAB. For each video frame, image features are tracked and a measure of algorithm confidence is computed. If the confidence value falls below a threshold, the tracked features and any necessary corrections are manually verified. The tracking algorithm relies on a passive black and white marker near the instrument tip and background lighting to improve instrument visibility and contrast. Using tracked features from both videos and calibration parameters of the stereo camera rig, instrument motion information is computed. Calibration parameters are found before data acquisition using the Camera Calibration Toolbox (Jean-Yves Bouguet, Camera Calibration Toolbox for Matlab, July 2010; http://www.vision.caltech.edu/bouguetj/calib_doc) for MATLAB.

Instrument motion is quantified using a three-dimensional position signal $\mathbf{P}(n)$ and three-dimensional orientation signal $\mathbf{O}(n)$ (Fig. 1). For both, n is an index representing video frame number. $\mathbf{P}(n)$ is the position of a point near the instrument tip in three-dimensional space over the course of captured video. $\mathbf{O}(n)$ is a three-dimensional unit vector whose origin is $\mathbf{P}(n)$. This vector is aligned with the axis of the rod connecting the instrument handle and tip. These signals represent instrument movement with 5 *df*: three for position and two for orientation. V-PITS cannot measure rotation about the instrument's axis. Measurement accuracy of the system was characterized as 0.15-mm root-mean-square using a precision mechanical stage. The system can measure motion over $3 \times 3 \times 3$ cm.

Phonomicrosurgical simulation station—A laryngeal dissection station used in phonosurgical education was employed.^{21,22} It simulates aspects of actual operations, including field-operating magnification, field of view, lighting, instrument movement freedom, and seating. Simulated surgical exercises are performed using cold phonomicrosurgery forceps (Pilling 50–6950) and scissors (Pilling 50–6818) inserted through a laryngoscope. Tasks are set at realistic operating distances and visualized with a Zeiss OPMI 1-H operating microscope (Carl Zeiss, Oberkochen, Germany) and objective lens with focal length of 38 cm. A modified mounting mechanism with a hollow tubular bracket 22 mm in diameter, simulating the trachea, allows for consistent placement of paper tasks rather than excised larynges. Paper tasks are secured to a square bracket attached to the mounting tube (Fig. 2).

Simulated surgical task—A 39×37.5 -mm piece of paper was printed with a grid of 3×3 -mm squares and a centered circle 28 mm in diameter. The circle serves as a guide for the laryngoscope field of view, which simulates glottis visualization. Within the circle, there are two red rectangles, three squares long. The rectangles are attached at the right vertical edge and free at all others; a small black dot is located center and above the top red rectangle (Fig. 3). The task required subjects to vertically cut at the right gridline of each red square,

grasping the paper with Babcock forceps in their nondominant hand and excising each red square with microscissors in their dominant hand. Each task required three cuts.

Subjects—Ten subjects (six male, four female; mean age, 20.7 ± 1.9 years) with no prior surgical experience (two medical students, eight undergraduate students) participated in the experiment, which was approved by the University of Wisconsin Health Sciences institutional review board. Subjects completed four surgical tasks a day at approximately the same time for 5 consecutive days. Data were collected for the first and last trial each day. Subjects were blinded to what trials included data collection. Before beginning the study, all subjects were informed of its purpose and instructed on task completion. Subjects could adjust the microscope and chair, and began each cut resting their instruments on the black dot. Two-minute breaks were provided between trials. Only the top red rectangle was used for each task.

Data Analysis

Motion metrics—For each task's video, six surges, or subtasks, were identified and tracked. Each surge consisted of a subset of frames over which the subject performed a single technique. The surges identified were grab 1, cut 1, grab 2, cut 2, grab 3, and cut 3. For “grab” surges, tracking data were generated for the forceps. For “cut” surges, tracking data were generated for the forceps and scissors. Instrument motion was measured for each surge. Three motion metrics were then computed (path length, depth perception, and motion smoothness) using the instrument motion signals $P(n)$ and $O(n)$. These metrics were chosen based on Cotin et al. (2002).¹⁶ Motion metrics are given by M_1 – M_3 in Equations 1, 3, and 4. Let $n = 0$ to $n = N-1$ be time samples from a subset of a recorded surgical exercise. Path length (M_1 , Equation 1) is the length of the curve covered by the instrument tip over time. This parameter describes the spatial distribution of the instrument tip in the workspace of the task.

$$M_1 = \sum_{n=1}^{N-1} \|P(n) - P(n-1)\| \quad (1)$$

Motion smoothness (M_2 , Equation 3) is derived along three axes and calculated using the k^{th} finite difference of a scalar discrete signal, given by Equation 2.

$$\nabla^k[f](n) = \sum_{i=0}^k (-1)^i \binom{k}{i} f(n-i) \quad (2)$$

$$M_2 = \frac{1}{N} \sqrt{\frac{1}{2} \sum_{n=0}^{N-3} (\nabla^3[x](n))^2 + (\nabla^3[y](n))^2 + (\nabla^3[z](n))^2} \quad (3)$$

Depth perception (M_3 , Equation 4) is the distance travelled along the instrument's axis. It quantifies depth-related position adjustment.

$$M_3 = \sum_{n=1}^{N-1} |(\mathbf{P}(n) - \mathbf{P}(n-1))^T \mathbf{O}(n)| \quad (4)$$

Task outcome—Cut precision was ranked from 1 to 3: 1 = excellent, clean cut down middle of line; 2 = fair, line cut mostly down middle, but not neatly; and 3 = poor, line missed or incomplete cut/tear. Six cuts total were scored for each subject each day to evaluate if increased practice correlated with increased surgical performance.

Statistical analysis—Motion metrics and task outcome were averaged for each day and subject. To determine if significant change over time was observed for motion metrics and task outcome, a one-way repeated measures analysis of variance was performed. To explore the relationship between task outcome and motion metrics, while accounting for individual variation, another one-way repeated measures analysis of variance was performed. If data did not fulfill parametric testing assumptions, a Friedman repeated measures analysis of variance on ranks was completed. A significance level of $\alpha = .05$ was used.

RESULTS

Path length, depth perception, and task outcome all showed improvement and displayed significant decreases in scores over time ($P < .001$, $P < .001$, $P < .001$, respectively). Motion smoothness did not display significant change over time ($P = .231$). Statistically significant relationships were observed between task outcome, and path length ($P < .001$), depth perception ($P < .001$), and motion smoothness ($P < .001$). For all scores, a lower value indicates a better score. Summary data are shown in Table I and Figures 4 through 7.

DISCUSSION

We simulated a phonosurgical task and used V-PITS to quantitatively evaluate surgical performance, selecting path length, depth perception, and motion smoothness as indicators. We expected increased practice time, and thus purportedly improved surgical performance, to negatively correlate with task outcome and motion metrics.

The significant decrease in task outcome scores demonstrates that as practice time increases, surgical performance improves. This supports the idea that training on simulators increases surgeon skill and technical competence.^{17,22} If study duration were longer, a plateau effect may have been observed, as seen in previous studies.²³ Given sufficient practice time, performance may reach a maximum.

The significant decreases in path length and depth perception scores over time are consistent with our hypothesis that as surgical performance improves, path length and depth perception scores improve. Accordingly, these motion metrics may be important indicators of phonosurgical performance. The significant relationships between improved surgical performance measured via task outcome and motion metrics further support the ability of V-PITS to evaluate phonosurgical performance. Although motion smoothness did not change significantly with time, consistent patterns within each individual may explain why motion smoothness exhibited a significant relationship with task outcome. The

nonsignificant change over time for motion smoothness is possibly due to small sample size and short study duration. Although metrics must be evaluated for this subspecialty independent of other surgical fields, previous findings indicate that motion smoothness is important in endoscopic and microsurgery.^{5,16} Given the similarities between these fields and phonomicrosurgery, it is probable that motion smoothness is significant for phonomicrosurgery. This motion metric may require longer practice time to improve; thus, longer study duration may reveal progress. The high standard deviations observed are likely due to variation in natural skill level between subjects and movement types. For example, greater path length is expected while positioning the forceps than while cutting. Analyzing surges may reveal lower standard deviations and give insight into which surges are most indicative of surgical performance.

As this project was the first step in developing V-PITS, the current system does not fully address the obstacles of expense and availability; these are necessary future optimizations. Several aspects of the system have potential to decrease costs. The software used is custom and not yet commercially available, but it constitutes the majority of the system's work. It runs on a desktop PC, eliminating the need for special processing equipment. The system was primarily composed of components readily available from major optical component suppliers. Future costs may also be reduced if operating room cameras are used, removing the expense of additional cameras.

One review article of systems used for tracking minimally invasive surgical instruments notes that many manufacturers and researchers do not report system accuracy.²⁴ Similarly, we found few reports of accuracy in the literature. Reported accuracies included $<40\text{ }\mu\text{m}$, $0.06\text{ in.} \pm 0.5\text{ mm}$, and 2.54 mm .²⁴ Although V-PITS does not exceed the best accuracy reported, it is higher than three of the four reported; furthermore, V-PITS is tailored toward phonosurgery.

This study was limited by low sample size and short duration. Accordingly, future studies will employ more subjects and extend beyond 5 days. The scoring system for task outcome was another limitation. Upcoming studies will use quantitative and objective scoring. Evaluation of deviation from the marked cut line shows promise for use as a scoring system. Error could be evaluated by counting pixels that diverge from an ideal cut area, giving a quantitative measure of accuracy; greater deviation from the ideal cut would correspond to a worse outcome. When possible, V-PITS will be used in conjunction with a global rating scale as instrument motion is only one factor that comprises surgical skill. However, eliminating this need in the future would save time and money, and obviate expert assessors. Additional data collection and motion metrics, such as number of movements (important in endoscopic and microsurgery^{5,7}), may be beneficial. Prospective studies will investigate using V-PITS to discriminate between skill levels and evaluate the effect of secondary parameters, such as arm supports, on surgical performance.

CONCLUSION

Our results suggest that instrument motion data collected by V-PITS may be useful for deriving scores of surgical performance, such as path length and depth perception.

Additional research is needed to understand the role of motion smoothness and capability of V-PITS to collect position data from which motion metrics can be derived. Further research may improve training and operating environments for phonomicrosurgery and credentialing of surgeons, thus creating better patient quality of care and reducing the impact of voice disorders.

Acknowledgments

The authors thank Glen Levenson, PhD, for his help in statistical design, Benjamin Smith, BS, for his assistance in engineering V-PITS parts, and Carlos Jaramillo, MSc, for his help in data analysis.

This study was funded by NIH grant number R01 DC008153 from the National Institute on Deafness and Other Communicative Disorders.

BIBLIOGRAPHY

1. Grantcharov TP, Bardram L, Funch-Jensen P, Rosenberg J. Assessment of technical surgical skills. *Eur J Surg*. 2002; 168:139–44. [PubMed: 12182237]
2. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ*. 2003; 327:1032–1037. [PubMed: 14593041]
3. Syme-Grant J, White PS, McAleer JPG. Measuring competence in endoscopic sinus surgery. *Surgeon*. 2008; 6:37–44. [PubMed: 18318087]
4. Richstone L, Schwartz MJ, Seideman C, Cadeddu J, Marshall S, Kavoussi LR. Eye metrics as an objective assessment of surgical skill. *Ann Surg*. 2010; 252:177–182. [PubMed: 20562602]
5. Reiley CE, Lin HC, Yuh DD, Hager GD. Review of methods for objective surgical skill evaluation. *Surg Endosc*. 2011; 25:356–366. [PubMed: 20607563]
6. Jensen AR, Wright AS, Levy AE, et al. Acquiring basic surgical skills: is a faculty mentor really needed? *Am J Surg*. 2009; 197:82–88. [PubMed: 19101249]
7. Saleh GM, Voyazis Y, Hance J, Ratnasothy J, Darzi A. Evaluating surgical dexterity during corneal suturing. *Arch Ophthalmol*. 2006; 124:1263–1266. [PubMed: 16966621]
8. Datta V, Mackay S, Mandalia M, Ara Darzi. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *J Am Coll Surg*. 2001; 193:479–485. [PubMed: 11708503]
9. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A. Relationship between skill and outcome in the laboratory-based model. *Surgery*. 2002; 131:318–323. [PubMed: 11894037]
10. Saleh G, Gauba V, Sim D, Lindfield D, Borhani M, Ghousayni S. Motion analysis as a tool for the evaluation of oculoplastic surgical skill. *Arch Ophthalmol*. 2008; 126:213–216. [PubMed: 18268212]
11. Grober E, Roberts M, Shin E, Mahdi M, Bacal V. Intraoperative assessment of technical skills on live patients using economy of hand motion: Establishing learning curves of surgical competence. *Am J Surg*. 2010; 199:81–85. [PubMed: 20103070]
12. Uemura M, Tomikawa M, Kumashiro R, et al. Analysis of hand motion differentiates expert and novice surgeons. *J Surg Res*. 2014; 188:8–13. [PubMed: 24418518]
13. Ezra D, Aggarwal R, Michaelides M, et al. Skills acquisition and assessment after a microsurgical skills course for ophthalmology residents. *Ophthalmology*. 2009; 116:257–262. [PubMed: 19091411]
14. Rolls AE, Riga CV, Bicknell CD, et al. A pilot study of video-motion analysis in endovascular surgery: development of real-time discriminatory skill metrics. *Eur J Vasc Endovasc Surg*. 2013; 45:509–515. [PubMed: 23465454]
15. Ramachandran S, Ghanem A, Myers A. Assessment of microsurgery competency—where are we now? *Microsurgery*. 2013; 33:406–415. [PubMed: 23712917]
16. Cotin, S.; Styopoulos, N.; Ottensmeyer, M.; Neumann, P.; Rattner, D.; Dawson, S. Metrics for laparoscopic skills trainers: the weakest link!. In: Dohi, T.; Kikinis, R., editors. *Medical Image*

Computing and Computer-Assisted Intervention—MICCAI 2002, Lecture notes in computer science. Vol. 2488. Heidelberg, Germany: Springer-Verlag; 2002. p. 35-43.

17. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg*. 1998; 175:482–487. [PubMed: 9645777]
18. Chen T, Dailey SH, Naze SA, Jiang JJ. The head-mounted microscope: A potential precision improvement in phonomicrosurgery. *Laryngoscope*. 2011; 122:781–784. [PubMed: 22434680]
19. Galleano R, Carter F, Brown S, Frank T, Cuschieri A. Can armrests improve comfort and task performance in laparoscopic surgery? *Ann Surg*. 2006; 243:329–333. [PubMed: 16495696]
20. Armstrong WB, Karamzadeh AM, Crumley RL, Kelley TF, Jackson RP, Wong BJF. A novel laryngoscope instrument stabilizer for operative microlaryngoscopy. *Otolaryngol Head Neck Surg*. 2005; 132:471–477. [PubMed: 15746864]
21. Dailey SH, Kobler JB, Zeitels SM. A laryngeal dissection station: educational paradigms in phonosurgery. *Laryngoscope*. 2004; 114:878–882. [PubMed: 15126748]
22. Verma SP, Dailey SH, McMurray JS, Jiang JJ, McCulloch TM. Implementation of a program for surgical education in laryngology. *Laryngoscope*. 2010; 120:2241–2246. [PubMed: 20939097]
23. Lee HH, Song KY, Park CH, Jeon HM. Training of surgical endoscopists in Korea: assessment of the learning curve using a cumulative sum model. *J Surg Educ*. 2012; 69:559–563. [PubMed: 22677598]
24. Chamarra MK, Grimbergen CA, Dankelman J. Systems for tracking minimally invasive surgical instruments. *Minim Invasive Ther Allied Techn*. 2007; 16:328–340.

Fig. 1. Signals used to quantify instrument motion. n = video frame number; $P(n)$ = three-dimensional position signal; $O(n)$ = three-dimensional orientation signal.

Fig. 2. Schematic of modified laryngeal dissection station equipped with the Video-Based Phonomicrosurgery Instrument Tracking System and paper surgical task.

Fig. 3. Simulated surgical task. Black rectangular area represents the precut area, the red area represents the target area, the black dot represents the starting point for each cut, and the triangular yellow shadow represents the projection of the far end of the laryngoscope. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

Fig. 4. Average path length scores for each day. Error bars represent standard error. A lower value indicates a better score.

Fig. 5. Average depth perception scores for each day. Error bars represent standard error. A lower value indicates a better score.

Fig. 6. Average motion smoothness scores for each day. Error bars represent standard error. A lower value indicates a better score.

Fig. 7.
Average task outcome scores for each day. Error bars represent standard error. A lower value indicates a better score.

TABLE I

Summary Data for Each Motion Metric and Task Outcome.

Practice Day	Path Length	Depth Perception	Motion Smoothness	Task Outcome
1	227 ± 115	112 ± 56	0.0257 ± 0.016	2.45 ± 0.37
2	107 ± 41	55 ± 24	0.0270 ± 0.012	2.10 ± 0.45
3	84.6 ± 52	46 ± 30	0.0288 ± 0.011	1.90 ± 0.37
4	68.1 ± 29	36 ± 14	0.0294 ± 0.012	1.70 ± 0.41
5	67.5 ± 20	34 ± 10	0.0294 ± 0.012	1.50 ± 0.32
<i>P</i> value	<.001	<.001	.231	<.001

Values are presented as mean ± standard deviation. A lower value indicates a better score. *P* values represent results of the repeated analysis of variance for change over time.