Construct validity of anterior segment anti-tremor and forceps surgical simulator training modules

Attending versus resident surgeon performance

Michael A. Mahr, MD, David O. Hodge, MS

PURPOSE: To compare the performance of the anterior segment forceps and anti-tremor training modules of the EYESi surgical simulator (VRmagic) by residents and experienced attending surgeons using the simulator for the first time.

SETTING: Department of Ophthalmology, Mayo Clinic, Rochester, Minnesota, USA.

METHODS: Twelve residents (4 per year) in the Mayo Clinic ophthalmology residency program and 3 experienced anterior segment surgeons participated. Each participant completed a total of 20 task trials on the EYESi forceps and anti-tremor level 4 training modules. Thus, the 15 participants completed a total of 300 task trials.

RESULTS: For the forceps module, experienced surgeons achieved significantly better total scores ($P = .03$), with lower total task time ($P = .007$) and instrument-in-eye time ($P = .006$) measurements. For the anti-tremor module, experienced surgeons achieved significantly better total scores ($P = .02$), with lower task time ($P = .04$) and instrument-in-eye time ($P = .02$) measurements. In addition, experienced surgeons performing the anti-tremor task had 76% more precise surgical outcomes as measured by the out-of-tolerance percentage ($P = .03$). All forceps and anti-tremor-measured parameters indicated significantly lower performance ($P<.05$) for the first 1 or 2 trials, with the exception of anti-tremor module incision stress, out-of-tolerance percentage, and average tremor values. Experienced surgeons had more consistent (lower variance) total scores on the forceps ($P = .05$) and anti-tremor ($P = .03$) training modules.

CONCLUSION: The EYESi surgical simulator anterior segment forceps and anti-tremor modules showed significant ($P<.05$) construct validity.

J Cataract Refract Surg 2008; 34:980–985 © 2008 ASCRS and ESCRS

The EYESi surgical simulator (VRmagic) was originally developed as a vitreoretinal surgical training device. The posterior segment membrane-peeling task has demonstrated construct validity, with experienced surgeons having fewer errors than novice surgeons.[1] Recent hardware and software advances have expanded the simulator to include anterior segment training tasks. These include forceps, anti-tremor, capsulorhexis, and phacoemulsification training modules.

Ophthalmology residency programs continue to seek better tools and methods for teaching, assessing, and documenting resident surgical competency in as safe and efficient a manner as possible.[2–4] Residency and fellowship training programs take great interest in the potential for commercially available virtual reality surgical simulators to further these aims.[5]

The EYESi simulator allows for the repeat performance and measurement of instructor-defined standardized surgical tasks. Its hardware and software measure and permanently record metrics that provide feedback in the following 4 main categories: Surgeon efficiency, achievement of surgical target or goal, surgeon error/tissue injury, and formative education/feedback during a task (EYESi Ophthalmic Surgery Simulator User Guide. Mannheim, Germany, VRmagic, 2006; 58–59).

To better understand the potential usefulness of the EYESi anterior segment forceps and anti-tremor training modules in training and assessing surgical

Published by Elsevier Inc.

0886-3350/08/$—see front matter
doi:10.1016/j.jcrs.2008.02.015

competency, we sought to evaluate the construct validity of these tasks.[6,7]

SUBJECTS AND METHOD

This study was performed using the EYESi surgical simulator (software version 2.2), a commercially available, proprietary hardware and software platform designed to simulate intraocular surgery (Figure 1). It includes a virtual operating microscope that provides a stereoscopic image to the trainee and requires that the user focus, pan, and zoom with a foot-pedal control. The EYESi also includes appropriately sized and shaped handheld probes that can virtually emulate various surgical instruments as well as anterior and posterior segment "heads" with model eyes that appropriately pivot and rotate when manipulated by the surgeon. Hardware and software continuously track the position of the instrument tips in the eye and superimpose this on the virtual intraocular environment. An instructor station and panel allow for real-time monitoring of performance, video replay and recording of recent tasks, manipulation of instrument settings (ie, phaco machine or vitrector), and viewing and downloading of individual and group historical performance data.

The study participants included all 12 Mayo Clinic ophthalmology residents (4 from each level of training) and 3 experienced anterior segment surgeons. Institutional review board approval was obtained, and all participants provided verbal consent.

All trials for each participant were performed during a single simulator session. With the exception of limited past "demos" of the EYESi for purchase consideration, this was each subject's first session using the EYESi forceps and anti-tremor task modules.

The EYESi anterior segment forceps module requires that the surgeon grasp 6 objects from the peripheral anterior chamber and place them in a "basket" in the center of the anterior chamber (Figure 2). Main outcome measures include time to complete task, iatrogenic corneal or lens injury, incision stress, and percentage of the task successfully completed. The real-world outcome that the forceps module attempts to train surgeons for is the ability to precisely grab the capsulorhexis flap while keeping the eye centered and not injuring the cornea or lens.

The EYESi anterior segment anti-tremor module requires the surgeon to precisely and efficiently move a small ball along a circular path on the surface of the anterior capsule (Figure 3). Good performance requires maintaining a circular path within relatively strict performance margins (submillimeter deviation from the intended path) while minimizing incision stress and lens/corneal injury. The real-world outcome that the anti-tremor task attempts to train and evaluate surgeons for is the ability to precisely pivot at the incision and control the instrument tip to create a curvilinear capsulorhexis while keeping the eye centered and not injuring the cornea or lens. For the trials in this study, the anti-tremor task required that the ball always be moved in a clockwise direction.

The study format was standardized for all subjects, and a single investigator (M.A.M.) supervised all tasks. Each participant was instructed to avoid caffeine use for 12 hours

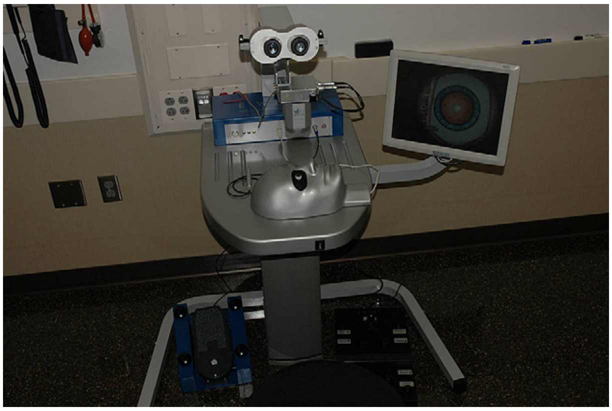

Figure 1. EYESi surgical simulator. Note left multifunctional phaco and vitrectomy foot pedal, right microscope pedal, and right touch screen instructor control and viewing display. The virtual operating microscope has a stereoscopic view with interchangeable "cataract head" hardware, shown here on the adjustable platform.

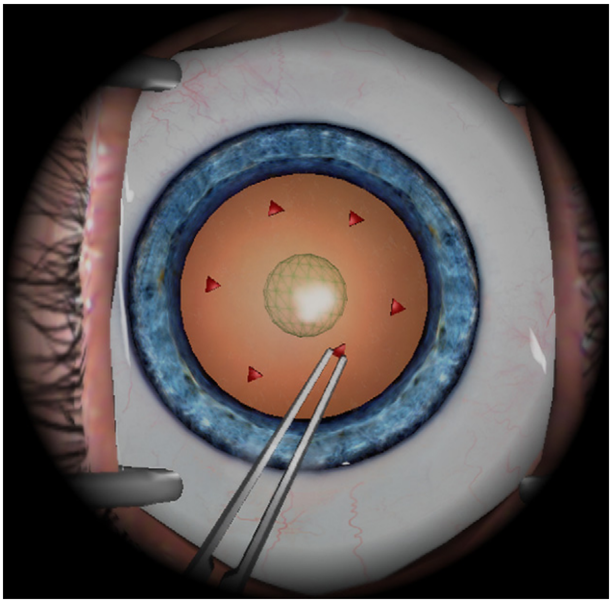

Figure 2. EYESi surgical simulator anterior segment forceps module. The surgical goal is to use the virtual forceps to move the 6 blocks from the peripheral anterior chamber to the mesh basket in the center of the anterior chamber.

Accepted for publication February 1, 2008.

From the Department of Ophthalmology, Mayo Clinic, Rochester, Minnesota, USA.

Neither author has a financial or proprietary interest in any material or method mentioned.

Presented as a poster at the World Ophthalmology Congress, Hong Kong, People's Republic of China, June 2008.

Supported by Research to Prevent Blindness, Mayo Foundation.

Corresponding author: Michael Mahr, MD, Department of Ophthalmology, Mayo W7, Mayo Clinic, 200 First Street SW, Rochester, Minnesota 55905, USA. E-mail: mahr.michael@mayo.edu.

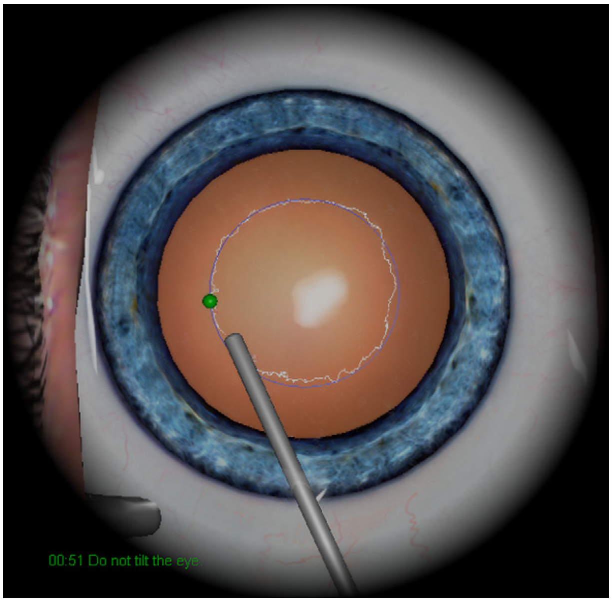

Figure 3. EYESi surgical simulator anterior segment anti-tremor module. The surgical goal is to precisely move the ball 360 degrees clockwise along the circular path (*blue ring*). At the end of the task, the surgeon's performance is graphically depicted by the white virtual "bread crumb trail." A sample intraoperative formative feedback message, "Do not tilt the eye," is seen in the lower left corner.

before the study session. In addition, residents were instructed that their performance would not be used as part of their residency training evaluation. The following information was recorded for each participant: age, sex, and hand dominance. All trials were performed in March 2007.

To begin, each enrollee received a basic orientation in machine setup including microscope adjustment, seating and positioning, and foot-pedal use. Each participant then performed 1 session on the EYESi anterior segment forceps level 1 (easiest level) and anti-tremor level 1 training tasks to get oriented to and familiar with the simulator hardware and virtual environment. The scores of the orientation session were not recorded, and no participant requested or was offered additional familiarization or training time.

The formal study trial then started. Each participant repeated 5 cycles in each trial of the following 4 sequential anterior segment task modules: right-hand forceps (level 4), left-hand forceps (level 4), right-hand anti-tremor (level 4), and left-hand anti-tremor (level 4). Thus, each subject performed 20 tasks: 5 forceps right hand, 5 forceps left hand, 5 anti-tremor right hand, and 5 anti-tremor left hand. The sequential order of trials alternated between hands, and tasks remained standardized and unchanged for all participants.

Total possible scores for each task ranged from 0 to 100 points. The EYESi awarded positive points for the percentage of each task goal successfully completed and subtracted points for efficiency and error parameters to arrive at a total score. The calculation for each task included point value deductions derived from the following measured parameters:

Forceps total score = number remaining objects (0 = +100 points) – excessive task time – incision stress – injured cornea area – lens displacement – number of out-of-focus instrument manipulations – forceps open on insertion or removal from the eye.

Anti-tremor total score = percentage of circle completed (100% = +100 points) – excessive task time – incision stress – injured cornea area – lens displacement – number of out-of-focus instrument manipulations – percentage of circle out-of-tolerance boundaries.

Statistical Analysis

For the forceps and anti-tremor modules, each directly measured and scored variable was analyzed for differences. A 3-factor repeated-measures analysis of variance with repeated measures on 2 factors was completed to evaluate the comparisons. The 3 factors included level of experience (resident versus attending), trial number (1 to 5 for each hand), and hand performing task (left hand versus right hand as well as hand dominance). Comparisons were adjusted for multiple comparisons using the Student-Newman-Kuels procedure. For both the forceps and anti-tremor modules, there were too few out-of-focus manipulation and corneal or lens-injury events to allow adequate statistical comparison.

The results by trial number (lower number = earlier trial) were compared across all groups to determine whether trial number was associated with performance. The comparisons were performed using the Student-Newman-Kuels procedure to adjust for the multiple comparisons.

The Student-Newman-Kuels procedure was also used to adjust for the multiple comparisons of hand and hand dominance. All parameters were compared for significant ($P < .05$) differences based on hand (left or right) and hand dominance (dominant versus nondominant).

As a measure of experienced surgeon versus resident performance consistency, a comparison of individual total score variances was performed for the forceps and anti-tremor modules. This was accomplished by calculating each participant's individual variance and performing a rank sum test.

Results are given as means $\pm$ SD.

RESULTS

Twelve residents (4 per training year) and 3 anterior segment surgeons were enrolled. Table 1 shows the participants' characteristics.

Experience Level

Table 2 shows the results for the forceps and anti-tremor modules based on level of experience (resident

Table 1. Resident and experienced surgeon characteristics.

Characteristic	Residents	Experienced Surgeons
Number	12	3
Sex, n (%)		
Male	6 (50)	3 (100)
Female	6 (50)	0
Age (y)		
Mean	30.0	44.7
Range	28–31	37–51
Dominant hand, n (%)		
Right	9 (75)	2 (67)
Left	3 (25)	1 (33)

Table 2. Trainee (resident) versus experienced (attending) surgeon EYESI anterior segment forceps and anti-tremor module level 4 performance.

	Forceps Module Trials			Anti-tremor Module Trials		
	Mean ± SD			Mean ± SD		
Parameter	Residents (n = 120)	Attending Surgeons (n = 30)	*P* Value	Resident (n = 120)	Attending Surgeons (n = 30)	*P* Value
Total score (points)	89.6 ± 9.52	97.5 ± 3.7	.03*	55.7 ± 30.6	88.6 ± 10.4	.02*
Total task time (min:s)	1:39 ± 0:32	0:58 ± 0:12	.007*	1:19 ± 0:20	1:03 ± 0:13	.04*
Time points lost (points)	−3.3 ± 1.8	−1.0 ± 0.7	.008*	−2.4 ± 0.8	−1.8 ± 0.6	.07
Time instrument inserted in eye (min:s)	1:22 ± 0:29	0:42 ± 0:08	.006*	1:02 ± 0:18	0:46 ± 0:09	.02*
Incision stress value	2.3 ± 3.5	0.4 ± 0.7	.12	1.8 ± 4.0	0.1 ± 0.2	.17
Incision stress points lost	−5.9 ± 7.3	−0.9 ± 2.4	.08	−4.4 ± 7.1	−0.1 ± 0.3	.05
Out-of-tolerance percentage value	NA	NA	NA	19.9 ± 16.0	4.7 ± 5.2	.03*
Out-of-tolerance percentage points, points	NA	NA	NA	−38.9 ± 29.8	−9.5 ± 10.5	.03*
Average tremor value	NA	NA	NA	36.6 ± 3.9	37.5 ± 5.1	.69

NA = not applicable
*Statistically significant difference ($P < .05$)

versus experienced surgeon). All participants completed 10 tasks for each module (forceps and anti-tremor); 300 trials were performed (120 resident and 30 experienced surgeon trials for each module).

Experienced surgeons achieved statistically significantly greater total scores (forceps $P = .03$ and anti-tremor $P = .02$) and were more time-efficient (lower total task and "instrument in eye" times) on the forceps and anti-tremor modules (*P* values ranging from 0.04 to 0.006). In addition, experienced surgeons were more precise and had fewer errors on the anti-tremor module, as evidenced by a significantly lower out-of-tolerance percentage ($P = .03$). There were no significant differences in incision stress or average tremor values.

Trial Number

For the forceps module, several statistically significant differences were observed. For total score (points), trial 1 (85.4 ± 12.4) was significantly different from each of the other trials ($P = .02$). Trials 2 (89.5 ± 9.0), 3 (92.8 ± 7.8), 4 (94.2 ± 6.2), and 5 (93.8 ± 6.9) were not significantly different from each other. For total time (minutes:seconds), trial 1 (1:58 ± 0:46) was significantly different from trials 2 (1:34 ± 0:29), 3 (1:27 ± 0:25), 4 (1:20 ± 0:21), and 5 (1:14 ± 0:18) ($P < .001$). Trial 2 was also significantly different from trials 4 and 5. In addition, trial 3 was significantly different from trial 5. For the time the instrument was inserted in the eye (minutes:seconds), trial 1 (1:37 ± 0:43) and trial 2 (1:19 ± 0:29) were significantly different from each other and all other trials. The times in trials 3 (1:09 ± 0:24), 4 (1:05 ± 0:20), and 5 (0:59 ± 0:17) were not significantly different ($P = .09$).

For the anti-tremor module, statistically significant trial differences were observed. For total time (minutes:seconds), trial 1 (1:29 ± 0:23) was significantly different from trials 2 (1:18 ± 0:23), 3 (1:13 ± 0:17), 4 (1:09 ± 0:13), and 5 (1:10 ± 0:16) ($P = .01$). No significant difference in anti-tremor module incision stress, out-of-tolerance percentage, or mean tremor value based on trial number was observed.

Hand and Hand Dominance

For the forceps and anti-tremor modules, the only statistically significant difference was in the mean tremor value on the anti-tremor module. Specifically, a comparison of the 75 right-hand versus 75 left-hand trials showed a significantly greater tremor value (ie, greater tremor measured) when the task was performed with the left hand (37.7 ± 3.8) than with the right hand (35.9 ± 4.3) ($P < .001$). When comparing hand dominance, there was no significant difference in average tremor value between dominant hands (36.5 ± 4.1) and nondominant hands (37.1 ± 4.2) ($P = .72$).

Experienced Surgeon Versus Resident Performance Consistency

Table 3 shows performance consistency based on a comparison of individual total score variances for the forceps and anti-tremor modules. For both

Table 3. Resident (attending) versus experienced surgeon EYESI total score variance.

Module	Mean Variance for Total Score		Rank Sum Test P Value
	Residents	Attendings	
Forceps	71.2	9.9	.05*
Anti-tremor	598.0	79.9	.03*

*Statistically significant

modules, the residents had significantly higher total score variances (less consistent outcomes) than the experienced surgeons (forceps $P = .05$, anti-tremor $P = .03$).

DISCUSSION

The EYESi Ophthalmic Surgery Simulator anterior segment forceps and anti-tremor modules showed significant ($P < .05$) construct validity when comparing experienced ophthalmology surgeons and residents.

The forceps module measurement and scoring parameters showed that experienced surgeons achieved outcomes 41% (total task time) to 49% (instrument time in eye) more efficiently than residents. The error measurements (corneal injury, lens injury, operating out-of-focus) occurred too infrequently to allow adequate statistical analysis. There was no significant difference in incision stress between groups (forceps $P = .12$ and anti-tremor $P = .17$), although all results trended toward higher incision stress for resident surgeons.

The forceps module serves as a robust task for the novice surgeon to train to navigate and perform manipulations in the anterior chamber, with the goal of maximizing efficiency while avoiding infrequent error events (corneal and lens injury). It currently does not have a scoring parameter that allows fine discrimination in measuring and evaluating more precise intraocular movements.

The anti-tremor module measurement and scoring parameters showed that experienced surgeons achieved 76% more precise outcomes (lower out-of-tolerance percentage value), 20% (total task time) to 26% (instrument time in eye) more efficient than the residents' outcomes. The anti-tremor module trains and measures the precision of intraocular movement in a rigorous manner. Its design makes it especially relevant in developing the skills necessary to perform a capsulorhexis. The anti-tremor module also places the surgeon in the realistic situation in which speed may negatively impact precision of movement and surgical outcomes.

Our standardized sequential trial design allowed for the determination of performance curves. It is apparent that most forceps module performance parameters reach a short-term plateau after the first 2 trials. For the anti-tremor module, the plateau may be less pronounced. These data support the notion that more formal performance assessments, using both the forceps and anti-tremor modules, may allow the subject to perform at least 3 trials, with the option of discarding results from the first 2 trials if the goal is to measure performance potential. Alternatively, to assess an individual surgeon's maximum performance (flat portion of performance curve), using the best 3 of 5 trials would be a reasonable strategy.

Our study found no significant differences ($P < .05$) in performance by hand (left or right) or hand dominance across all parameters with the following exception: The mean tremor values for the anti-tremor module were statistically significantly greater for left hands than for right hands ($P < .001$), although there was no significant difference based on hand dominance. Although we cannot adequately explain this observation, it may be related to the fact that the anti-tremor trial design required that both right-hand and left-hand tasks navigate the "ball" in a clockwise direction. Performing a future trial with the direction changed to counterclockwise may provide further insight. Anecdotally, most participants were initially hesitant to perform tasks with the nondominant hand because of apprehensions about poor performance. By the end of the trials, most were favorably surprised by their high performance when working with the nondominant hand. Although most surgeons do not regularly perform precise intraocular maneuvers (ie, capsulorhexis) with the nondominant hand, our data suggest that doing so might result in reasonable performance and efficiency outcomes for both novice and experienced surgeons. An alternative consideration includes the possibility that using the nondominant hand requires a greater level of cognitive effort or concentration that provides a similar technical outcome but "wastes" the surgeon's efforts, which could be better applied to other cognitive operative issues. In a study by Kageyama et al.,[8] resident surgeons attained unexpectedly good outcomes performing phacoemulsification using their nondominant hand. Their findings may support our interpretation. In addition, it is possible that our study was not adequately powered to demonstrate subtle but clinically significant differences in performance based on hand dominance.

The difference between less experienced and more advanced surgeons often goes beyond the ability to adequately perform a surgical task or reach the desired surgical goal. Instead, the expert is more likely to achieve performance targets in a reproducible manner in which sequential operations have little variation in outcomes. Both the forceps and anti-tremor module total score variances show that experienced surgeons were significantly more likely to achieve better outcomes with less variability (forceps $P = .05$, anti-tremor $P = .03$). Specifically, the experienced surgeons' mean total score standard deviations were 63% lower than those for the residents on both the forceps and anti-tremor modules. Future assessments of surgical performance should consider lower variability in outcome measures as an important measure of surgeon competency.

REFERENCES

1. Rossi JV, Verma D, Fujii GY, Lakhanpal RR, Wu SL, Humayun MS, de Juan E Jr. Virtual vitreoretinal surgical simulator as a training tool. Retina 2004; 24:231–236
2. O'Day DM. Assessing surgical competence in ophthalmology training programs [editorial]. Arch Ophthalmol 2007; 125:395–396
3. Binenbaum G, Volpe NJ. Ophthalmology resident surgical competency; a national survey. Ophthalmology 2006; 113: 1237–1244
4. Lee A. Teaching and assessing surgical competency in ophthalmology [letter]. Ophthalmology 2006; 113:2380–2381
5. Khalifa YM, Bogorad D, Gibson V, Peifer J, Nussbaum J. Virtual reality in ophthalmology training. Surv Ophthalmol 2006; 51:259–273
6. Downing SM. Validity: on meaningful interpretation of assessment data. Med Educ 2003; 37:830–837
7. Dunn WF. Simulators in Critical Care and Beyond. Des Plaines, IL, Society of Critical Care Medicine, 2004; 126–127
8. Kageyama T, Yaguchi S, Metori Y, Chida M, Koizumi K, Onishi T, Ayaki M. Visual results and complications of temporal incision phacoemulsification performed with the non-dominant left hand by junior ophthalmologists. Br J Ophthalmol 2002; 86: 1222–1224