*Medicine Meets Virtual Reality 2002*
*J.D. Westwood et al. (Eds)*
*IOS Press, 2002*

# Reliability and Validity of EndoTower, a Virtual Reality Trainer for Angled Endoscope Navigation

Randy S. HALUCK MD FACS[1], Anthony G. GALLAGHER PhD[2,3],
Richard M. SATAVA MD FACS[2], Roger WEBSTER PhD[4],
Thomas L. BASS MD[1], Cynthia A. MILLER BS[1]
[1]*Department of Surgery, Penn State College of Medicine, Hershey, PA 17033*
[2]*Department of Surgery, Yale University School of Medicine, New Haven, CT 06510*
[3]*School of Psychology, The Queen's University of Belfast, Belfast BT7 1NN, NI*
[4]*Department of Computer Science, Millersville University of Pennsylvania School of Science and Mathematics, Millersville, PA 17551*

**Abstract**. We hypothesized that a simulator designed to train surgical novices angled laparoscopic navigation would show an improvement in subjects' performance, a high test re-test reliability and high internal validity as measured by standardized coefficient alpha. It was also predicted that simulator performance would be strongly related to objectively assess perceptual and visuospatial ability. EndoTower has good face validity in that it mimics precisely the performance of an angled laparoscope and previous studies suggest construct validity. In this study, EndoTower is shown to be a trainer intended for laparoscopic skill than can test for inherent perceptual and visuospatial ability.

## 1. Background/Problem:

The Penn State EndoTower (EndoTower) is a PC based simulator/trainer that utilizes the Virtual Laparoscopic Interface (VLI, Immersion Corp., San Jose, CA). An additional hardware component attached to the VLI device represents an angled laparoscope structure and function. It is designed to provide instruction and practice in the skills necessary for camera/laparoscope navigation for minimally invasive surgery [1].

All surgical procedures require some degree of visual navigation within an operative field. Laparoscopy presents additional challenges as visual and spatial cues are altered. Most commercially available angled laparoscopic camera and lens systems must be independently rotated, or not rotated by the user in order to achieve the desired angle of view while maintaining correct right-side-up orientation. This camera and angled rigid lens system is used in many different operations covering a number of surgical specialties.

The function, orientation, and correct manipulation of an angled laparoscopic lens in combination with an independently rotating camera system is not necessarily intuitive. As the surgeon performs the actual operation, the duty of camera navigation falls to the assistant, often a less-experienced team member. Robotic assisted camera navigation also requires that the surgeon have a fluid understanding of the workings of the optical system.

Several studies suggest that an inherent trait, visuospatial ability, correlate with a higher level of surgical skill [2-4]. Visuospatial abilities relate to the ability to mentally

represent three-dimensional situation using key landmarks, and to have a clear mental picture of their relationship in space [5]. While other studies do not support this conclusion [6], it is apparent that some degree of visuospatial ability is required for surgical performance,

Surgeons using video assistance need to form visual impressions of three-dimensional structures, organs and instruments, from a two-dimensional television monitor. While this is often described as loss of binocularity [7], it is simpler and more accurate to call it pictorial perception. So-called primary cues – binocular disparity and convergence, accommodation, and motion parallax – are present in abundance. The difficulty is that the primary cues and other cues related to lighting and texture yield a conclusion that is inimical to surgery. The conclusion specified is that the structures in view form a single surface, virtually flat and usually vertical. The surgeon has to set aside that conclusion in order to register the information carried by subtler 'pictorial' cues and to reconstruct the structure that they specify despite the incompleteness of the information provided. Individuals differ in this ability and such differences could clearly contribute to performance differences for pictorially guided laparoscopic surgery.

Although EndoTower was designed to provide instruction in angled laparoscope navigation, we hypothesized that it would also serve as a testing platform to reliably assess visuospatial and perceptual ability. A simulation tool that is able to reliably assess these abilities, identifying levels of inherent ability in surgeons, while increasing skill acquisition may have additional value for surgical education.

## 2. Methods:

The task for the EndoTower testing session entailed identifying randomly placed arrows in the virtual "EndoTower". The EndoTower is a three-dimensional (3D) block tower with cylindrical holes within its various arms (Figure 1). The EndoTower provides a relatively complex three-dimensional structure for exploration with the virtual angled lens. In the simulator, 3D arrows of varying colors and directional orientation are randomly placed around the tower and inside the cylindrical cut-outs. Arrows were chosen as the targets so that participants would be required to maintain right-side-up orientation in order to be able to properly identify them. Arrows could have one of four colors and six directional orientations giving 24 possible combinations. Six arrows were randomly placed on and within the virtual EndoTower as part of the software program (Figure 2).

To operate the simulator, the subject held one of the VLI laparoscopic tools, modified to simulate the angled lens laparoscope. Rotation of this mock camera and lens rotated the horizon and angle of view from the longitudinal axis of the virtual laparoscope, respectively. The virtual world mimics the real world as the user's viewpoint is placed at the tip of the laparoscopic tool. Due to the 3D nature of the EndoTower and the cylindrical cutouts, identifying all the arrows entailed manipulating the laparoscope to many different and challenging orientations and positions.

Once a target arrow was found and identified, the student selected the matching arrow from a list using the foot-pedal function of the VLI. Any collision with the EndoTower caused the view to become fuzzy and blurry with 'red out', simulating touching the laparoscope to an organ and smudging the lens with blood. The learner was required to withdraw the virtual laparoscope and await a cleaning mode, thus penalizing the user's score in terms of time efficiency and also scoring an error. EndoTower mechanism recorded the amount of time to correctly identify the color and orientation of each arrow and the number of arrows.

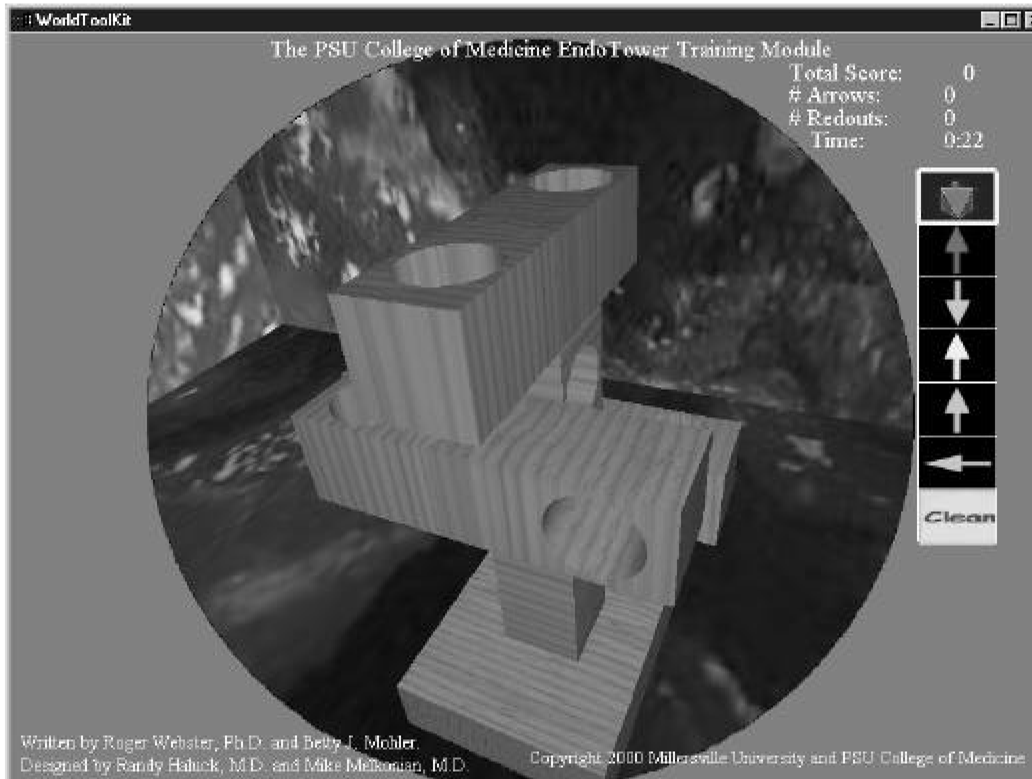


Figure 1: The three dimensional EndoTower

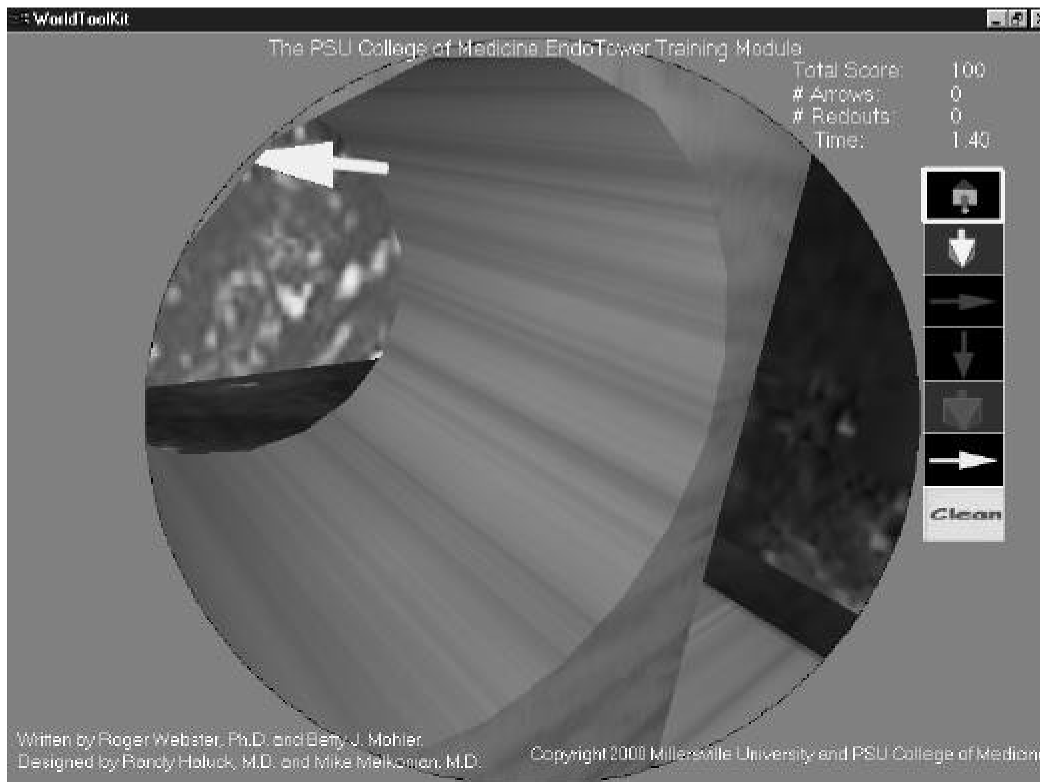


Figure 2: A randomly placed arrow within the virtual EndoTower

A time limit of four minutes was selected to complete the task of correctly identifying all six arrows. Computerized scoring was determined by the time to find all six of the target arrows plus the number of "red-out" errors multiplied by a factor of 10. If at the end of the four minute time limit, all of the arrows were not identified, the number of arrows not found was multiplied by 30 and added to the score. Therefore, a lower score indicated a better performance.

Twenty-five laparoscopic novices completed three trials on EndoTower. None of the subjects had ever used EndoTower or manipulated a rigid laparoscopic lens. All subjects received a brief and standardized instruction session on the goals and function of EndoTower. One testing was accomplished and the computer-generated score was recorded.

All subjects also completed the Pictorial Surface Orientation (PicSOr) test battery, a perceptual task that assesses the ability to reconstruct 3-D from 2-D perceptual cues [8]. They also completed three tests that assess fundamental visuospatial abilities, Card Rotation (CR), Cube Comparison (CC), and Map Plan (MP) [9]. Gallagher et al have reported that these tests are reliably associated with laparoscopic performance [10, 11]. The study was approved by the institutional review board with informed consent obtained from all participants.

## 3. Results:

The mean EndoTower scores were, trial 1 = 337 ± 50, trial 2 = 324 ± 73 and trial 3 285 ± 71. Subjects' performance improved consistently across the three trials. Analysis of variance for repeated measures showed that there was a statistically significant difference between the trials ($F(2, 24) = 13.05$, $p < 0.00001$). Differences between the different trials were compared for significance with Scheffe Post Hoc F-Tests. Although there was an improvement in scores between trial 1 and trial 2 this improvement was not statistically significant. The improvement in scores between trial 1 and 3 was statistically significant (Scheffe $F = 12.62$, $p < 0.0001$) as were the differences between trials 2 and 3 (Scheffe $F = 6.87$, $p < 0.01$). A statistically significant within subject effect was also observed. The within subject effect means that some subjects' EndoTower scores were statistically different from the mean EndoTower score ($F(2, 24) = 5.05$, $p < 0.0001$).

The test re-test reliability of EndoTower was calculated using Pearson's Product Moment Correlation Coefficient. The observed correlation coefficient was $r = 0.8$ ($F(1, 23) = 33.82$, $p < 0.0001$) and $r^2 = 0.6$.

To determine the internal validity or consistency of EndoTower as a measurement device Standardize Coefficient Alpha was calculated and was observed to be $\propto = 0.88$ ($F(2, 24) = 13.574$, $p < 0.00001$).

The correlations between EndoTower performance and PicSOr, CR, CC and MP were $r=0.591$, $r=0.296$, $r=0.354$ and $r=0.392$ respectively. However, when any single visuospatial test was combined with the PicSOr perceptual test in a multiple regression model the correlations increased dramatically. The results for the different regression models showed that for, PicSOr and CR, $r=0.658$, $r^2=0.43$ ($F(2,37)=14.2$; $p<0.00001$); PicSOr and CC, $r=0.701$, $r^2=0.49$ ($F(2,37)=17.89$; $p<0.00001$); and PicSOr and MP, $r=0.708$, $r^2=0.50$ ($F(2,37)=118.6$; $p<0.00001$).

## 4. Discussion / Conclusions:

A number of factors cumulate in the successful performance of an operation. Spencer suggested that "probably about 75% of the important events in an operation are related to making decisions, and about 25% to dexterity" [12]. Although this may be true, an operation is the result of manipulation of tissues ultimately requiring technical performance and skill. At this point in time, the acquisition of operative technical skill is poorly understood, both from the standpoint of the ingredients or abilities that serve as the basis for skill and which individuals might posses those ingredients. A test that replicates tasks for which surgical skill is required, yet can assess inherent abilities, may help to further the understanding of skill acquisition. One factor that will affect operative performance is practice.

The human brain has a limited capacity to attend to different sources and amounts of information. This means that the surgeon can only attend to a given amount of information at any given time. The more familiar the surgeon is with information, or the more frequently they have come in contact with it, the less attentional processing resources that information will require. The opposite is also true. This is important in surgery because attentional resources that are used on behaviors such as instrument navigation are not available for the most important aspect of surgery, i.e., surgical performance. We believe that simple but fundamental aspects of video assisted surgery such as camera navigation should be largely learned in the laboratory so that residents can hone their skills in the operating room rather than acquire them from basics.

One of the primary functions of a simulator, virtual reality or otherwise, is to improve performance. Another function of a virtual reality simulator is to provide objective assessment of performance, however, assessment cannot be accepted at face value. Satava [13] has suggested that if a virtual reality system purports to assess performance, validation metrics associated with that claim must be reported. In this paper we have reported that EndoTower has good face validity in that it mimics precisely the performance of an angled laparoscope. Interestingly, EndoTower also appears to be able to discriminate those individuals who have difficulty performing the task as indicated by the significant probability value for between subjects ANOVA results. It also has been demonstrated to have high test re-test reliability and a high alpha coefficient. In this study, we have also taken the validation a stage further.

Visuospatial ability and perceptual ability have been hypothesized and reported to be correlated with laparoscopic performance [10, 11]. These hypotheses appear sensible but need further experimental validation. In this study these abilities were found to be strongly related to EndoTower performance, particularly when included in a multiple regression model. In fact, simple regression models accounted for 43–50% of the variance. This is an important point because multimodal behaviors (e.g., perceptual, cognitive and psychomotor) required for complex real-world performance, such as surgery are unlikely to be explained by a simple correlation model, more likely a multiple regression model or ideally a Structural Equation Model.

Intuitively, it would seem that an ability to mentally reconstruct 3D from 2D cues and manipulate complex 3D structures and environments would be fundamental to performing operative manipulations. However, there is disagreement in the surgical literature regarding their ability to predict surgical skill. The study reported here shows that objectively assessed perceptual and visuospatial abilities predict angled scope navigation performance. This lends support to the hypothesis that these abilities are related to laparoscopic surgical performance. Furthermore, EndoTower assessment of these abilities has greater face validity for the assessment of a surgical skill, i.e., the test is

simulated angled scope navigation. However, further validation work on this issue is required.

Virtual reality simulation has moved from 'is it pretty' to 'does it work and where is the evidence' to sophisticated metrics about performance parameters. Satava [13] has suggested that all virtual reality simulators must meet these fundamental criterion markers to be considered as a scientifically validated simulator. EndoTower has gone some way to meet these criteria however one of the major challenges for virtual reality remains, "Does training in the virtual environment transfer to the OR?" This work is currently underway in our laboratories and we are optimistic about the outcome.

**5. Acknowledgements:**

This project was funded in part by the Penn State Eberly Virtual Hospital Project and one of the authors (AGG) was supported by the Fulbright Distinguished Scholar program.

**6. References:**

1. Haluck RS, Webster R, Melkonian M, Mohler B, Dise M, LeFever A: A virtual reality surgical training module for instruction in angled laparoscopic lens navigation. Stud Health Technol Inform 2001;81:171-176.
2. Schueneman AL, Pickleman J, Hesslein MA et al: Neuropsychological predictors of operative skill among general surgery residents. Surgery 1984;96:288-295.
3. Gibbons RD, Baker RJ, Skinner DB: Field articulation testing: A predictor of technical skills in surgical residents. J Surg Res 1986;41:53-57.
4. Risucci D, Tortolani A, Horowitz M: Neuropsychological assessment of general surgery interns. Focus Surg Edu 1991;10:14-15.
5. DesCoteaux JG, Leclere H: Learning surgical technical skills. CJS 1995;38:33-38.
6. Francis NK, Hanna GB, Cresswell AB, Carer FJ, Cuschieri A. The performance of master surgeons on standard aptitude testing. Am J Surg 2001;182:30-33.
7. Reinhardt-Rutland AH and Gallagher AG. Visual depth perception in Minimally Invasive Surgery. In SA Robertson (Ed) Contemporary Ergonomics;1996:531-536.
8. Cowie R. Measurement and modeling of perceived slant in surfaces represented by freely viewed line drawings. Perception 1998;27:505–540.
9. Ekstrom RB, French JW, Harman HH, Dermen D. Manual Kit for Factor – Referenced Cognitive Tests. 1976 Princeton New Jersey: Educational Testing Service.
10. Gallagher AG, Crothers I, Satava RM. Comprehensive objective assessment of fundamental skills for laparoscopic surgery. Surg Endosc 2001; In Press.
11. Gallagher AG, Cowie R, Crothers I, Jordan JA, McGlade K, McGuigan J, McClure N. PicSOr: An objective test of perceptual skill that predicts laparoscopic performance in three separate studies. Surg Endosc 2001; In Press.
12. Spencer FC. Observations on the teaching of operative technique. Bull Am Coll Surg 1983;3:3-6.
13. Satava, RM. The need for metrics in surgical education. (1999) Surg Endosc; 13: 1082-1082