Surg Endosc (2002) 16: 943–948
DOI: 10.1007/s00464-001-9146-y

SURGICAL ENDOSCOPY
and Other Interventional Techniques

The Advanced Dundee Endoscopic Psychomotor Tester (ADEPT) objectifying subjective psychomotor test performance

M. P. Schijven,[1] J. Jakimowicz,[1] C. Schot[2]

[1] Department of Surgery, Catharina Hospital Eindhoven, Michelangelolaan 2, P.O. Box 1350, 5602 ZA Eindhoven, The Netherlands
[2] Department of Clinical Physics, Catharina Hospital Eindhoven, Michelangelolaan 2, P.O. Box 1350, 5602 ZA Eindhoven, The Netherlands

Received 23 July 2001/Accepted in final form: 18 October 2001/Online publication: 5 March 2002

Abstract
Background: This study was undertaken to establish the value of the Advanced Dundee Psychomotor Tester (ADEPT) as an objective real-time scoring system, correcting for subjective assessor opinion on endoscopic task performance. The main research questions were as follows:

- Are surgeons good estimators of their own performance on ADEPT?
- Do surgeons perceive ADEPT to be a valid instrument for measuring laparoscopic skills?
- Does performance on ADEPT reflect innate psychomotor ability?

Methods: Each of 45 surgeons completed two runs on ADEPT. The runs comprised five standardized tasks. A posttest visual analog scaled (VAS) questionnaire measuring attitude toward skills testing in general, validation, and performance on ADEPT was used. Subjective responses were compared with objective scores generated through performance on ADEPT.
Results: Surgeons emphasize the importance of using a variety of training methods for surgical residents during their residency, including laparoscopic virtual reality simulators. Monitoring of residents' endoscopic progress seemed to be a key issue. Surgeons themselves underestimate their individual performance on ADEPT (mean subjective score of 6.1 vs mean objective score of 6.6). Self-reported performance on ADEPT is unreliable because confidence intervals between the VAS score and the ADEPT score overlap. Surgeons disagree on the validity of ADEPT. The mean score for validity was 5.8, ranging from 0 to 10 with almost equal distribution over the scale. Innate ability established as surgeons' scores express is high concordance between test run and true run, with 72.7% of the participants' true run score within one distance from the test run.
Conclusions: Surgeons cannot correctly predict their standardized individual test result on ADEPT. Performance on ADEPT reflects innate psychomotor ability along with improvement over runs. Surgeons are ambivalent in assessing the validity of ADEPT, irrespective of personal performance.

Key words: Surgical skills — Virtual reality — Simulation — Psychomotor testing — Objective assessment

Correspondence to: M. P. Schijven

To improve and evaluate endoscopic task performance, attempts have been made to establish objective means for assessment of performance. The Advanced Dundee Endoscopic Psychomotor Tester (ADEPT) was developed for this purpose in 1997 in the Ninewells Hospital at the University of Dundee [9]. In addition to evaluating endoscopic task performance, this device can be of value in evaluating innate psychomotor ability and training methods for minimal access surgery [15]. The research questions for our study were as follows:

- Are surgeons good estimators of their own performance on ADEPT?
- Do surgeons perceive ADEPT to be a valid instrument in measuring laparoscopic skills?
- Does performance on ADEPT reflect innate psychomotor ability?

Methods

Equipment

A computer-controlled device, ADEPT was developed for objective evaluation of endoscopic task performance (Fig. 12). Its hardware consists of a dual gimbal mechanism that accepts 4.8-mm standard endoscopic instruments for bilateral manipulation in a defined three-dimensional workspace. The device has three ports: one to mount a standard endoscope and two to mount manipulation instruments.

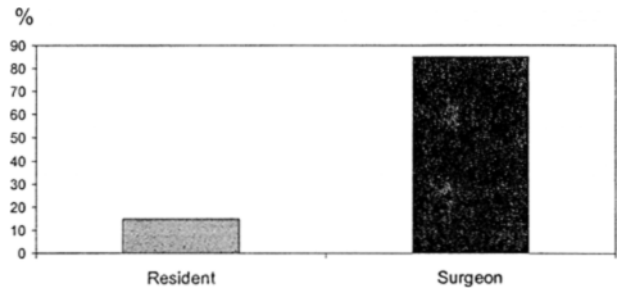

Fig. 1. Occupation.

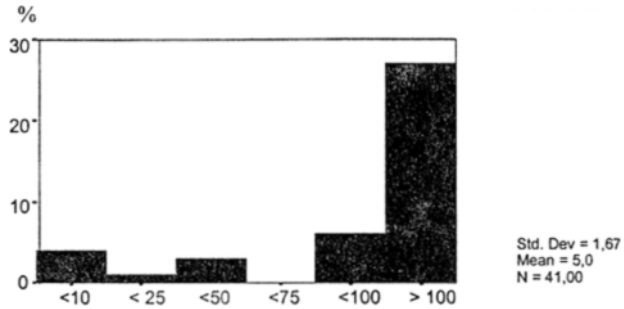

Fig. 2. Number of laparoscopic procedures.

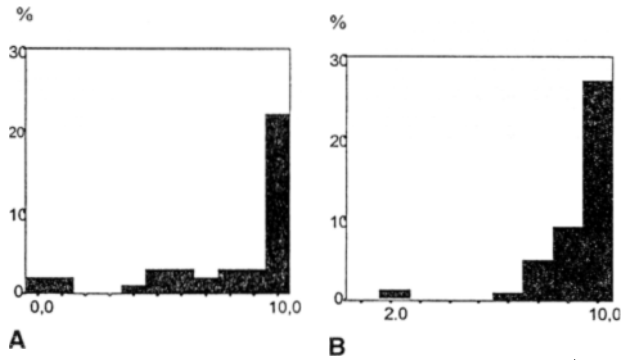

Fig. 3. **A** Importance of training by virtual reality. **B** Importance of training through basic surgical skills course.

On the isocenter of the device, a task box is placed. This box comprises five different target tasks based on the main actions involved in endoscopic manipulation. The tasks include movements such as grasping, moving, and positioning an object, for example, manipulating a top plate with one instrument while negotiating the underlying task with the other instrument. The five target tasks in the task box are overlaid by a transparent spring-mounted top plate with access apertures. The task box itself is mounted on a spring-mounted base plate. Contact between one of the instruments and the lined edges of any access aperture results in a so-called "probe error," simulating overload in pressure.

The five task targets in the box are a flick-switch, a rotation dial, a joystick, and two slider tasks positioned at different angles. A self-running computer program has been developed, which randomly assigns any number of specified tasks during a test session. The maximum time allocated to each task was 60 sec. Elapsed time during performance of target tasks; success in completion of task; horizontal, vertical, and rotational movements of the instruments; plate and probe errors are recorded through an interface unit, then translated into data by the computer.

For our study, ADEPT was set up to achieve optimal angles for endoscopic manipulation [10, 13]. The elevation angle of the endoscope was chosen so that the optical angle was perpendicular to the task box (task box horizontal plane, 25°; endoscope mount, 65°; distance to task box, 10 cm). Left and right instruments were each mounted at 35° so that elevation angles were 60°. The azimuth angle (between each grasper and endoscope) was 30°, and the manipulation angle (between left and right graspers) was 60°. The base–plate angle (distance between the two probe mounts) was 75°. A standard Storz cold light fountain 450-V light source with a standard Storz Endovision 9050-PB single-chip camera, a Sony high-resolution monitor, and a 0° forward-viewing 10-mm endoscope 33 cm in length were used (Karl Storz, Tuttlingen, Germany).

Participants

The subjects participating in this study were 45 congress participants in the 9th International Congress of the European Association of Endoscopic Surgery, held 13–16 June 2001, in Maastricht, The Netherlands. Their endoscopic surgical experience varied from that of a surgical resident with no laparoscopic experience to that of a fully trained consultant (attending) surgeon.

Outcome measures

Two runs of five randomly assigned tasks, each with a maximum allocated time period of 60 sec, were performed by each participant. Execution time, successful completion, total plate error time, and total probe error time per task were recorded. The number of tasks completed in one run without any plate error, without probe error, and without either plate or probe error was recorded as perfect plate task run, perfect probe task run, and perfect task run, respectively.

Subjective assessment

The participants' opinion was asked on a variety of subjects using a questionnaire with a 10-cm continuous response scale, a visual analog scale (VAS). Statements on the necessity for standardized laparoscopic and virtual reality training courses for surgical residents were proposed, as well as questions for performance on ADEPT. Endpoints on the VAS for the statements ranged from 0 cm (complete disagreement) to 10 cm (complete agreement). An escape answer "do not know" was available. Other statements related to personal performance on ADEPT, its validity for teaching purposes and assessment of laparoscopic skills.

Statistics

Frequency tables and box plots were constructed for visualization of data. Student's paired *T*-test was performed. The Statistical Package for the Social Sciences, version 9.0, was used.

Results

Demographics

The 45 participants originated from 16 countries. Their ages ranged from 27 to 61 years (mean 42.3 years). Of these participants, 80% were right-handed, 10% left-handed, and 10% ambidextrous. Their level of experience varied from that of resident to that of surgeon (Fig. 1), with 15% working as residents, 85% as surgeons. In terms of laparoscopic surgical experience, 19.5% of the participants performed fewer than 50 laparoscopic operations, 14.6% fewer than 100, and 65.9% more than 100 laparoscopic procedures per year (Fig. 2). Only one surgeon performed fewer than 100 laparoscopic procedures per year, and no resident performed more than 100 laparoscopic procedures per year.

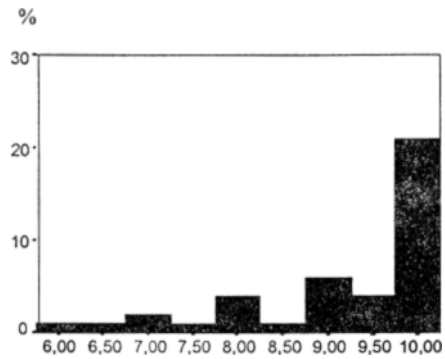

Fig. 4. Importance of monitoring laparoscopic skills.

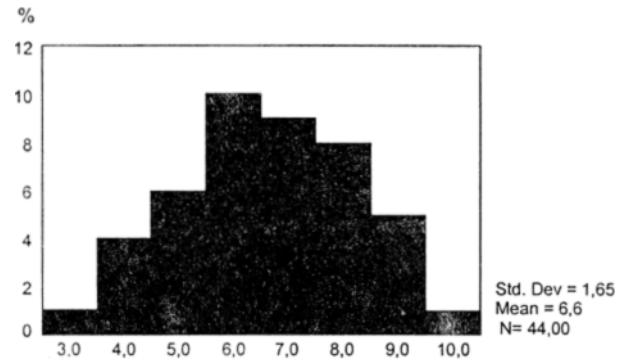

Fig. 6. Distribution of SUM scores.

Components:

-number of successful tasks second run:	0-5
-time run < 150 sec equals "6":	1
-number of "perfect tasks" ≥ 1	1
-number of "perfect probe tasks" ≥ 1	1
-number of "perfect plate tasks" ≥ 1	1
-test run ≠ 0 (at least one task successful)	1
	10

Fig. 5. SUM.

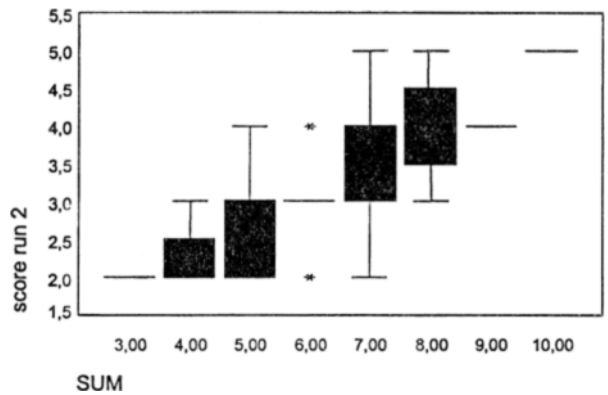

Fig. 7. Correlation between score and other components of the SUM score.

Questionnaire

Almost all participants agreed with the statement that it is necessary for surgical residents-in-training to participate in a (basic) laparoscopic skills course before operating on patients (mean 9.3 ± 1.34). Most of the participants thought it was important to train surgical residents-in-training on laparoscopic virtual reality simulators such as MIST-VR and Xitact before they operated on patients (mean 7.9 ± 3.04; Fig. 3). Furthermore, most of them considered an objective assessment method for monitoring the progress of residents' laparoscopic skills to be a valuable asset (mean 9.16 ± 1.12; Fig. 4).

Surgeons' perception of performance on ADEPT

For each participant, self-perceived performance, as reported on the VAS scale, was compared with the participant's score on ADEPT. A positive correlation between the two variables indicated that the surgeon was a reliable assessor of performance. Participants' score on ADEPT was reflected through the variable "SUM" (Fig. 5). This variable was computed as follows: successful tasks (0 to 5 points per run), total execution time of less than 150 sec (1 point), number of "perfect tasks" (1 point), number of "perfect probe tasks" (1 point), number of "perfect plate tasks" (1 point), and score on test run (1 point for at least one task successful).

The result from this calculation is a maximum SUM score of 10. The SUM distribution for the participants is normal in shape (Fig. 6), and the elements in SUM were in concordance (Fig. 7). Therefore, SUM seems to be a valid objective estimate of performance. The SUM score was plotted against reported performance scores resulting from the questionnaire (Fig. 8). The Mean VAS for the test run was 5.4 ± 3.8, and the mean VAS for the true (second) run was 6.1 ± 3.4. The mean score for "overall performance on ADEPT" was 6.1 ± 2.2, and the mean score on SUM was 6.6 ± 1.64. Thus, surgeons tend to underestimate performance on ADEPT slightly, but not significantly $\rho = 0.25$, paired T-test. Nonetheless, the VAS does not seem to be a reliable predictor of objective performance because the confidence intervals for the scores of SUM all seem to overlap. In other words, surgeons' SUM score cannot be predicted by knowing his or her self-reported score for overall task performance.

Surgeons' perception of ADEPT as a valid instrument for measuring laparoscopic skills

Most of the surgeons (78.6%) were not familiar with ADEPT. The participants who were acquainted with ADEPT (21.4%) had either heard of it or read published work on if. None of the participants had been exposed previously to ADEPT. The participants can thus be considered unbiased because no one had ever worked with ADEPT previously, and most had never heard of it. Scores for validity given by the participants averaged

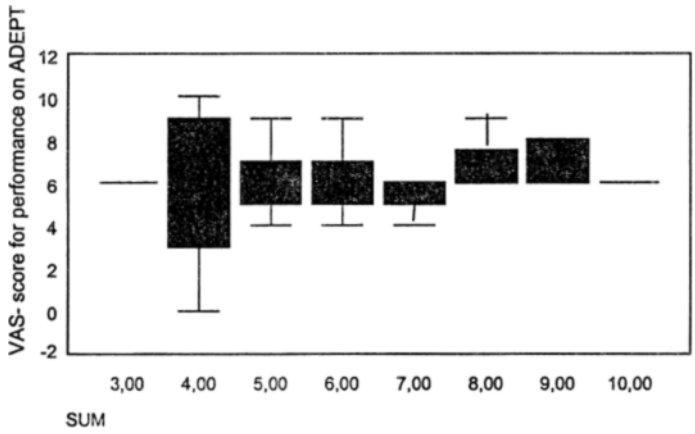

Fig. 8. Is the visual analog scale (VAS) score a reliable estimator of objective performance?

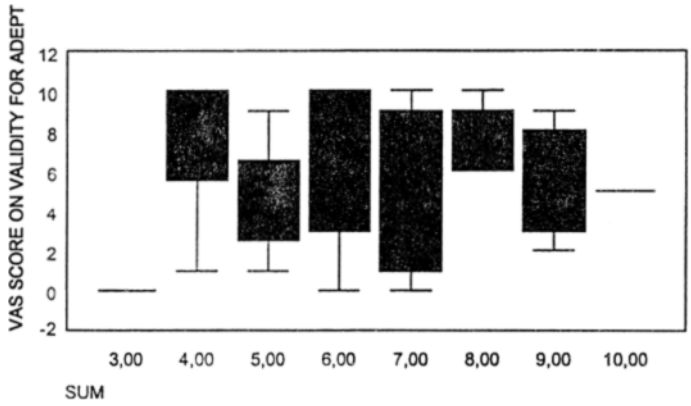

Fig. 9. Is ADEPT a valid instrument, corrected for individual performance?

5.8, with a range of 0 to 10 distributed almost equally on the VAS scale. It is interesting to see that opinion on validity, corrected for individual performance by SUM, was not uniform (Fig. 9). Individuals performing well on ADEPT did not rate its validity any higher than individuals performing poorly. Interestingly, inexperienced surgeons (who had performed fewer than laparoscopic procedures) were not worse in terms of SUM score than experienced surgeons (who had performed more than 100 laparoscopic procedures; Fig. 10). Therefore, ADEPT seems capable of excluding experience in laparoscopic surgery as a determinant of importance in assessing psychomotor ability.

Comparison of performance between test and true runs

The participants showed some improvement in total time needed for runs (test run mean time of 129 ± 54.4 sec vs true run mean time of 100 ± 39.3 sec). Also, the number for successful task performances on the true run was higher than for the test run (2.4 ± 1.3 vs 3.3 ± 1). However, innate ability was established because the surgeons' scores expressed a high concordance between test run and true run. More specifically, 72.7% of the participants expressed a true run score within one scoring distance from the test run. On paired-sample Student's *T*-test, both variables, time and score, differed significantly ($p = 0.001$ for both variables), indicating some improvement or training effect on ADEPT. When the data are plotted, (Fig. 11), it can be seen that, in general, the participants who performed poorly on the test run did not perform much better on their true run. Also, the participants who initially performed well on the test run were likely to perform well on the true run. Only one participant had a score of 0 on the test run and a score of 5 on the true run. This participant obviously used the test run as an exercise to get fully acquainted with ADEPT.

Discussion

The usual format for measuring surgical endoscopic skills uses subjective assessment methods of performance, with or without the use of structured rating scales. More objective methods of assessment are important for the validation of performance. Furthermore, objective measurement methods may identify a person's innate ability for endoscopic task performance [12, 15]. Finally, objective measurement methods can be useful in establishing concurrent validity for a variety of laparoscopic training settings, which can vary from the well-known

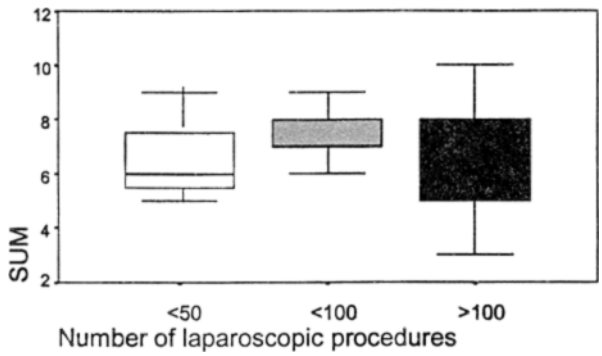

Fig. 10. Construct validity.

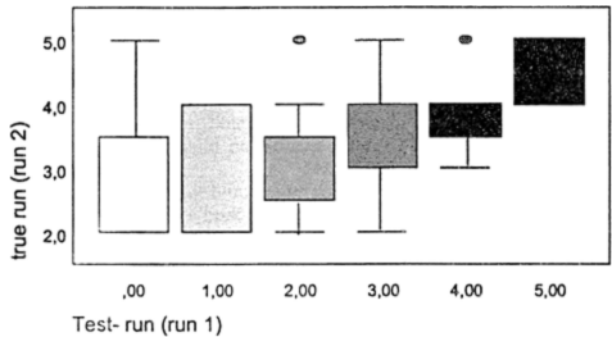

Fig. 11. Innate ability.

and cheap endoscopic training box or mannequin to more expensive laparoscopic skills training settings using animal material. Recently, skills training methods simulating laparoscopic procedures such as MIST-VR, LapSim, and Xitact, all computer controlled laparoscopic virtual reality trainers, have become the object of study [5, 14, 16]. Important abilities in endoscopic surgery are controlled precision, two-hand coordination, steadiness, and aiming [2, 11]. Psychomotor skill acquisition for laparoscopic surgery is a difficult but essential prerequisite for safe surgery. It must be noted, however, that many other important factors influencing surgical performance must be taken into account such as knowledge, personality traits, skills, fatigue, operating room staff surrounding the surgeon, laparoscopic instrumentation, and stress resulting from the actual performance of surgery [4]. By measuring psychomotor visual–spatial and perceptual abilities, ADEPT may be considered an important and feasible device for monitoring endoscopic skills assessment.

Our study used ADEPT to focus on concordance between subjective and objective assessments of psychomotor skills performance. Attempts have been made to develop an objective instrument for estimating surgical skills performance. Structured questionnaires for observation of surgical performance are used [18], as well more objective outcome denominators such as precision and speed for laparoscopic task performance [1]. However, there still is a subjective component in these scores because they are derived from human observation. Also, important aspects such as pressure and range of motion are not taken into account. In contrast,

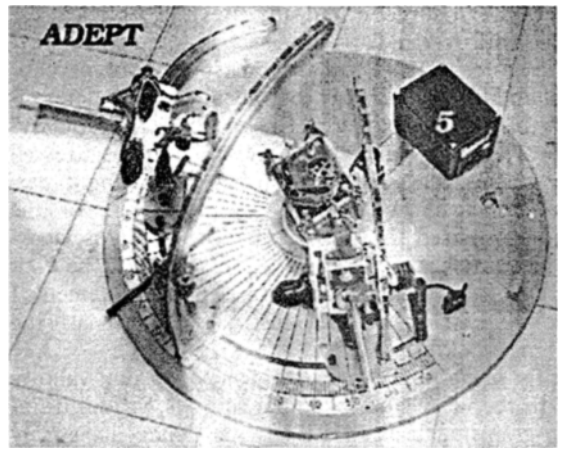

Fig. 12. ADEPT system. **1** Base plate. **2** Probe mounts. **3** Scope mount. **4** Task box. **5** Interface unit.

ADEPT, the device of our study, does take these aspects into account, along with elapsed time and successful task performance. Face validity and concurrent validity of ADEPT was ensured previously by Francis et al. [3] and Macmillan and Cushieri [15].

Our first research question was: Are surgeons good estimators of their own performance on ADEPT? Most studies compare observatory results with the outcome of task performance. Surgeons' estimate of performance seldom is compared with an objective outcome parameter of task performance. Our results show that, in fact, little trust can be placed in participants' own estimate of performance. Because estimation of performance seems to be so difficult, the need for an objective scoring system to evaluate endoscopic task performance is obvious.

Our second research question was: Do surgeons perceive ADEPT to be a valid instrument for measuring laparoscopic skills? This question refers primarily to the external validity of ADEPT, comparing task performance with the clinical laparoscopic setting. Before external validity can be assessed, internal validity of the instrument must be secured. Our study shows ADEPT's internal validity to be high. Participants were assessed alike performing a standardized test procedure. By establishing the stable and reliable indexed performance score SUM, as shown by this study, ADEPT can be regarded as a powerful objective scoring system for endoscopic psychomotor testing. For external validity, ADEPT has previously been shown to express a strong correlation with clinical competence [15]. However, in our study, surgeons did not rate validity of ADEPT highly. Moreover, the surgeons did not express uniformity in their opinion about ADEPT'S validity. Experienced surgeons did not perform worse or better on ADEPT than inexperienced surgeons. It is a fact that the endoscopic setting of ADEPT is not similar to the clinical laparoscopic situation. More precisely, ADEPT is a *psychomotor* tester, focusing on only one aspect of a complex area of interacting determinants (cognitions, skills, abilities) influencing laparoscopic surgery. This might explain why experienced surgeons do not perform

any better. No surgeon had any previous practical experience with ADEPT.

Therefore ADEPT indeed measures pure psychomotor ability without any of the additional determinants that make someone a good or experienced surgeon. The true purpose of ADEPT must be stressed before performance. Otherwise, frustration about performance may occur. Therefore, the nature of the question proposed to the participants actually is incorrect. The external validity of ADEPT would be estimated better by evaluating responses to the statement: I believe ADEPT is a valid instrument in measuring psychomotor skills.

As for using ADEPT to establish concurrent validity of laparoscopic (virtual) training settings, virtual training programs such as MIST-VR and Xitact have only recently been developed, and currently are being further refined and studied. Because surgeons must inevitably go through a learning curve in performing operations, in current practice, patients still have to pay the costs, not only literally speaking. The introduction of varied laparoscopic procedures calls for a more structured approach toward attainment of technical and ultimately clinical competence [8]. Computer-based virtual reality training programs, once validated, have the potential to solve many of the economic, educational, ethical, and safety issues related to the process of becoming a surgeon [6, 7]. The extent to which skills can be transferred from a skills training environment to the true clinical environment is likely to be dependent on the similarity of the setting. Advocates of surgical virtual reality simulator technology predict a revolution in surgical education. Although considerable enthusiasm for this concept exists, the development and implementation of simulation technology in surgical training has been limited by the absence of skills assessment devices [17]. The process and progress of teaching residents in a surgical skills laboratory environment for laparoscopic skills can be monitored by using ADEPT. For a variety of laparoscopic teaching programs, ADEPT may act as a reliability and validity check in addition to and in comparison with standard rating scales for performance.

Our third research question was: Does performance on ADEPT reflect innate ability? Macmillan and Cushieri [15] have shown ADEPT to be a system identifying aspects of performance that do not improve with practice. Indeed, in our small study analyzing two runs of five tasks, both runs had a high concordance of 72.7%. As discussed earlier and visualized by Figs. 10 and 11, scores on ADEPT seem to independent of laparoscopic experience, and scores are stable over runs. However, some training effect from the use of ADEPT itself seems to play a significant role because both time and number of successful tasks are significantly higher among the second run statistics.

References

1. Derossis AM, et al. (1998) Development of a model for training and evaluation of laparoscopic skills. Am J Surg 175: 482–487
2. Fleishman E (1996) Acquisition of skill, Academic Press, New York pp 147–167
3. Francis N, Hanna GB, Cushieri A (2001) Reliability of the Advanced Dundee Endoscopic Psychomotor Tester for bimanual tasks. Arch Surg 136: 40–43
4. Gallagher A, Crothers I, Satava R (2001) Comprehensive objective assessment of fundamental skills for laparoscopic surgery. In 9th International Congress of the European Association for Endoscopic Surgery, Maastricht, The Netherlands
5. Gantert WA, et al. (2001) The "virtual patient": a VR simulator for training in laparoscopic surgery: initial experience in a training course. In: 9th International Congress of the European Association for Endoscopic Surgery, Maastricht, The Netherlands
6. Grantcharov TP, et al. (2001) Virtual reality computer simulation: an objective method for the evaluation of laparoscopic surgical skills. Surg Endosc 15: 242–244
7. Haluck R, K, TM (2000) Computers and virtual reality for surgical education in 21st Century. Arch Surg 135: 786–792
8. Hamdorf JM, Hall JC (2001) Acquiring surgical skills. Br J Surg 87: 28–37
9. Hanna, GB, et al. (1998) Computer-controlled endoscopic performances assessment system. Surg Endosc 12: 997–1000
10. Hanna GB, Cushieri A (1999) Influence of the optical axis-to-target view angle on endoscopic task performance. Surg Endosc. 13: 371–375
11. Hanna GB, Drew T, Clinch P (1996) A microprocessor-controlled psychomotor tester for minimal access surgery. Surg Endosc 10: 965–969
12. Hanna GB, Drew T, Cushieri A (1997) Technology for psychomotor skills testing in endoscopic surgery. Semin Laparosc Surg 4: 120–124
13. Hanna GB, Shimi S, Cushieri A (1997) Influence of direction of view, target-to-endoscope distance, and manipulation angle on endoscopic knot tying. Br J Surg 84 1460–1464
14. Hyltander A, Rhodin P, Liljegren E (2001) Virtual reality training skills can be transferred to the operating room: a prospective and randomized study. In: 9th International Congress of the European Associating for Endoscopic Surgery, Maastricht, The Netherlands
15. Macmillan AI, Cushieri A (1999) Assessment of innate ability and skills for endoscopic manipulations by the Advanced Dundee Endoscopic Psychomotor Tester: predictive and concurrent validity. Am J Surg 177: 274–277
16. Satava M, GA, MJ (2001) Reliability and validity of the minimally invasive surgical trainer virtual reality system. In: 9th International Congress of the European Association for Endoscopic Surgery, Maastricht, The Netherlands
17. Smith CD, Stubbs J, HD (1998) Simulation technology in surgical education: can we assess manipulative skills and what does it mean to the learner. In: Medicine meets virtual reality: art, science, technology: healthcare (R) Evolution. In: Wesjtwood JD, et al. IOS Press and Ohmsa pp 379–380
18. Szalay D, et al. (2000) Using operative outcome to assess technical skill. Am J Surg 180: 234–237