Journal of Surgical Research **121,** 112–119 (2004)
doi:10.1016/j.jss.2004.02.005

How to Select Aspirant Laparoscopic Surgical Trainees: Establishing Concurrent Validity Comparing Xitact LS500 Index Performance Scores with Standardized Psychomotor Aptitude Test Battery Scores

Marlies P. Schijven, M.D., H.Sc.,*[,1] Jack J. Jakimowicz, M.D., Ph.D.,† and Fiona J. Carter, M.D., Ph.D.‡

**Department of Surgery, Erasmus Medical Center, Dr. Molewaterplein 40, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands; †Department of Surgery, Catharina Hospital Eindhoven, Michelangelolaan 2, P.O. Box 1350, 5602 ZA Eindhoven, The Netherlands; and ‡Department of Surgery and Molecular Oncology, Ninewells Teaching Hospital and Medical School, University of Dundee, DD1 9SY Dundee, Scotland, United Kingdom*

Submitted for publication January 16, 2004

Background. **Although a controversial topic in medical education, the selection of aspirant surgical trainees is a subject that needs to be addressed. In the view of preventing surgical trainee drop-outs and of appropriate allocation of limited resources, it is an issue critical to the profession. Traditional methods of selection are often subjective, and do not seem to correlate with skill needed for surgery. Standardized neuropsychometric test batteries may be useful in helping to select aspirant laparoscopic surgeons. Our study attempts to link surgical novices' psychometric ability test battery data with actual performance outcome on an objective, validated, and reproducible surgical laparoscopic task using virtual-reality simulation.**

Materials and methods. **Thirty-three novices with no laparoscopic surgical experience participated. Each participant performed the Xitact LS500 Virtual Reality cholecystectomy clip-and-cut module 30 times. Individual learning curves were computed and patterns were assessed. Participants were examined using the aptitude test battery including the Abstract Reasoning test, the Space Relations test, the Gibson Spiral Maze test, and the Crawford Small Parts Dexterity tester.**

Results. **Over 900 virtual-reality simulation tasks were generated and assessed. Of the participants, 93.3% were able to complete the virtual-reality simulation 30 times and all elements of our aptitude test battery. The abstract reasoning test is the only test correlating significantly to Xitact test outcome. This test is highly correlated to the space relations test. Both tests have discriminative power, comparing groups of performers.**

Conclusions. **The present study addresses the concurrent validity in aptitude testing, comparing scores of surgical novices on the Xitact LS500 laparoscopic cholecystectomy virtual-reality simulation with performance scores on a battery of standardized psychometric aptitude tests. The abstract reasoning and the space-relation test have predictive and selective value, identifying individuals who have good laparoscopic surgical virtual-reality performance. Aspiring laparoscopic surgeons who score below 25 on either test, that is, an expected 36%, would have to be further assessed using Xitact surgical task performance. The group of participants scoring above 35 on the Abstract Reasoning test and above 45 on the Space Relations test, that is, an expected 18% of the population, is unlikely to mal-perform on Xitact. The other 46% could very well benefit from Xitact simulation and assessment when the opportunity is present.**

Key Words: **virtual reality; minimal access surgery; laparoscopy; simulation; learning curve; validation; selection.**

INTRODUCTION

In aviation, pilot skills have, obviously, a large impact on flight safety. A pilot needs certain basic (innate) abilities which cannot be trained, but which are assessed in the aspirant pilots. Selected pilots thus possess a certain set of abilities that acts as the framework for aviation skills acquisition through training. Studies show that, next to measurements of cognition,

[1] To whom correspondence and reprint requests should be addressed at Department of Surgery, Erasmus Medical Center, Dr. Molewaterplein 40, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands. Fax: +31 10 4635307. E-mail: mschijv@zonnet.nl.

0022-4804/04 $30.00

psychomotor tests are valid selection tools and have prognostic significance in pilot training programs [1, 2]. In parallel, surgical skill is clearly linked to the clinical outcome of surgical procedures. Preliminary studies have indeed shown a correlation between psychomotor test systems used in aviation and the endoscopic performance of surgical trainees [3].

Although controversial, an issue that needs to be addressed in view of appropriate allocation of limited resources is the selection of surgical trainees. Medical educators tend to pay much attention to the design of the curriculum, but little to the selection of students [4]. Effective and fair selection of aspirant surgeons is under heated debate, but nevertheless, critical to the profession. Traditional measures such as Medical College Admission tests, National Board Scores, and CGE levels do not correlate [5] or even correlate negatively [6] with subsequent surgical performance ratings. Ratings, which are subjective in themselves because they are derived from a senior surgeons' subjective assessment, use unstructured observational methods. However, this is still the predominantly used method of assessment of surgical skill.

Neuropsychologic and neuropsychometric tests, assessing visuospatial perception, stress tolerance, and psychomotor ability, have been reported to be valid and consistent predictors of future surgical performance in multiple studies [6–9]. Nevertheless, attempts to establish a *standardized* battery of aptitude measures has proved unsuccessful, indicating inconsistency and counterintuitive findings between scores on different tests [10, 11]. The main factors accounting for the variability in those studies were the lack of objectivity in assessing participants' surgical task performance and the use of only one surgical procedure for the assessment of task performance, that is, no longitudinal studies have assessed surgical task performance data. It is clear that a proper identification of predictive psychomotor abilities is needed before the issue of selection of surgical trainees can even be discussed. Thus, one needs to establish a valid aptitude and psychomotor ability test battery that has potential for predicting future surgical performance. One way of doing so is by linking surgical novices' psychometric ability test data on various aptitude tests to their outcomes on a validated, reproducible surgical task. Nowadays, virtual-reality medical simulation offers, next to excellent standardized training opportunities, the possibility of unbiased, structured, validated, and repetitive assessment of surgical task performance. The Xitact LS500 is such a virtual-reality laparoscopic simulator. This simulator mimics, among others, the cholecystectomy clip-and-cut procedure. The simulation is thought to integrate psychomotor, visuospatial, as well as cognitive elements associated with the clip-and-cut procedure of the laparoscopic cholecystectomy.

The present study addresses the concurrent validity in aptitude testing, comparing scores of surgical novices on the Xitact LS500 laparoscopic cholecystectomy virtual-reality simulation with performance scores on a battery of standardized psychometric aptitude tests.

MATERIALS AND METHODS

Subjects

Only surgical novices were selected as participants to measure pure abilities and to exclude individuals with skills derived from previous surgical experience. Thirty-three hospital residents and final-year interns—unconditioned for laparoscopy—participated in the study. Participants received a 1-h familiarization protocol on the Xitact simulator, introducing them to the laparoscopic cholecystectomy clip-and-cut scenery. Participants followed a step-by-step teaching schedule, consisting of live video clips, a color-guided teaching approach with instruction on common errors and problems, and a free-format clip-and-cut exercise. Feedback through Xitact's assessment sheet and the instructor were given after this 1-h familiarization. Each participant then performed the clip-and-cut exercise 30 times, that is, 10 times per session for three consecutive days. Learning curves were then computed for each participant [12]. Based on the individual learning curve pattern, participants were classified into one of four groups of performers. One month later, the participants took a battery of standard aptitude tests.

Simulator

The Xitact® LS500 laparoscopic cholecystectomy simulator is a modular virtual-reality training platform, developed for training and education of a variety of laparoscopic skills (Fig. 1). It is a hybrid simulator, combining a physical object (The OpTable, or "virtual abdomen") with a computer software simulation, providing the visual image and tactile feedback. The Xitact incorporates Basic Surgical Skills, the Clip-and-Cut task of the laparoscopic cholecystectomy, a Camera Navigation module, and a Peritoneal Dissection module for dissection of Calot's triangle. The module used for determining learning curves is the Clip-and-Cut task. This module has been previously validated [13, 14]. The Xitact LS500 is developed and registered by Xitact SA, Morges, Switzerland (Fig. 1).

Xitact Performance Groups

Previous research on the Xitact LS500, assessing over **900** simulation task runs of **30** laparoscopic novices, revealed that 16.7% of participants are naturally gifted and did not need much simulation training to achieve adequate task proficiency. The majority of participants (63.3%) were able to achieve laparoscopic task proficiency over 30 simulation runs. Nevertheless, 20% of the participants seemed to be lacking innate manipulative abilities to such an extent that they could not achieve a stable task performance [12]. Based on these results, different Xitact performance group profiles were constructed.

Group 1 consisted of performers with a high level of innate abilities, gaining little extra improvement through VR training (16.7% of total group).

Group 2 consisted of performers with a moderate level of innate abilities, gaining improvement and stability through VR training (30% of total group).

Group 3 consisted of performers with a moderate level of innate abilities, gaining unstable improvement through VR training (33.3% of total group).

FIG. 1. The Xitact® LS500 laparoscopic cholecystectomy simulator.

Group 4 consisted of performers with a low level of innate abilities, not gaining improvement through VR training (20% of total group).

Aptitude Tests Battery

The Aptitude Test Battery included the Abstract Reasoning test and the Space Relations test, which are subtests of the Technical Abilities Battery of the Differential Aptitude Test (Psychological Corp. Ltd., London, UK). Also, the Gibson Spiral Maze was included in the Battery (Gibson HB, 1961: Hodder and Stoughton, London, UK), as well as the Crawford Small Parts Dexterity Tester (Psychological Corp. Ltd.). Each test is thought to be indicative of a certain psychometric trait or ability.

The Abstract Reasoning test investigates an individual's nonverbal reasoning ability and is related to IQ. Subjects must reason with geometric designs. They have to try to complete 40 tasks of abstract reasoning, choosing the right sequel option following a series of logical, abstract patterns within 20 min.

The Space Relations test investigates individual's visuospatial ability. Subjects must mentally reconstruct a 3D object from a 2D pattern and rotate this object in mind in space. They have to compete 50 tasks within 25 min. A multiple-choice standardized answering sheet for analysis of results is used for both tests.

The Gibson Spiral Maze test measures eye–hand coordination. Subjects must trace a line through a printed paper maze, avoiding obstacles. Standardized vocal stress-enhancing triggers are administered at interval bases during the test. Execution time and error score are end-point parameters.

The Crawford Small Parts Dexterity tester measures eye–hand coordination and manual dexterity. Part one involves placing small pins into holes on a metal plate, followed by fitting collars over the pins. Execution time is the end-point parameter.

Statistics

The Statistical Package for the Social Sciences SPSS version 9.0 was used for statistical calculations. A correlation matrix using either Pearson's or Kendall-tau's $_b$ correlation coefficient—whenever appropriate—was constructed. Also, groups were clustered and compared for significant differences in test outcome using the Mann–Whitney U test with the significance level set at 0.05.

RESULTS

Demographics

Mean age of the participants was 28 years, ranging from 21 to 35 years. There were 18 males and 15 females in the study, 19 were right-handed, 2 left-handed, and 2 ambidextrous. The participants were described as follows: 11 participants were interns; 2 residents were in training for emergency medicine; 6 residents were in training for radiology; 3 were in training for urology; 1 was in training for cardiology; 3 were in training for pulmonology; 2 were in training for anesthesiology; and 5 were in training for internal medicine. Three participants could not fully complete the required 30 runs. The learning curves were thus calculated for 30 participants. Of these participants, 93.3% were able to take part in the Aptitude Test Battery. Two participants could not take part, and they were therefore omitted from further analysis.

Correlation Matrices and Scatter Plots

The Aptitude Test Battery was compared with Xitact Group Classification in Correlation Matrix 1 (Table 1). Only the Abstract Reasoning Test seems to correlate with the Xitact Performance Groups. The Abstract Reasoning Test in itself is highly related to the Space Relations Test, revealing that abstract reasoning and visuospatial ability are highly related concepts. Also, Gibson's Spiral Maze time and error are correlated, depicted in Correlation Matrix 2 (Table 2).

Plots were constructed to visualize the correlations. Plot 1 shows a negative correlation between error and time for the Spiral Maze, indicating that the longer a participant takes to draw the spiral maze, the less error is likely to occur (Table 3). Plot 2 shows a positive correlation between the Abstract Reasoning Test and the Space Relation Test, indicating that individuals are consistent in scoring either high or low on both tests.

Discriminative Power of the Aptitude Test Battery

For this analysis, the better performers on Xitact (groups 1 and 2) were compared with the ones who performed badly (groups 3 and 4; Table 4). On comparison of the means of these two clusters, only the Ab-

TABLE 1

Correlation Matrix 1

Aptitude tests correlation matrix		Crawford Small Parts Dexterity Tester	Gibson Spiral Maze (time)	Gibson Spiral Maze (error)	Abstract Reasoning Test	Space Relations Test
GROUP						
N	28	27	28	28	28	28
Corr. Coeff.	1.000	0.003	0.134	−0.028	−0.310*	−0.260
Sig.		0.983[b]	0.364[b]	0.852[b]	0.038[b]	0.085[b]
Crawford Small Parts Dexterity Tester						
Corr. Coeff.	0.003	1.000	0.096	0.484*	−0.216	−0.229
Sig.	0.983[b]		0.633[a]	0.011[a]	0.278[a]	0.251[a]
Gibson Spiral Maze (time)						
Corr. Coeff.	0.134	0.096	1.000	−0.515**	−0.310	−0.065
Sig.	0.364[b]	0.633[a]		0.005[a]	0.108[a]	0.742[a]
Gibson Spiral Maze (error)						
Corr. Coeff.	−0.028	0.484*	−0.515**	1.000	0.131	−0.022
Sig.	0.852[b]	0.011[a]	0.005[a]		0.507[a]	0.910[a]
Abstract Reasoning Test						
Corr. Coeff.	−0.310*	−0.216	−0.310	0.131	1.000	0.845**
Sig.	0.038[b]	0.278[a]	0.108[a]	0.507[a]		0.000[a]
Space Relations Test						
Corr. Coeff.	−0.260	−0.229	−0.065	−0.022	.0845**	1.000
Sig.	0.085[b]	0.251[a]	0.742[a]	0.910[a]	0.000[a]	

[a] Correlation coefficient used is Pearson.
[b] Correlation coefficient used is Kendall's tau-$_b$; both variables were ranked.
* Correlation is significant at the 0.05 level (two-tailed).
** Correlation is significant at the 0.01 level (two-tailed).

stract Reasoning Test and the Space Relations Test proved to be discriminative.

Concurrent Validity

Table 5 shows a distribution plot of the concurrent validity between Xitact Performance Groups and the Aptitude Test Battery.

TABLE 2

Correlation Matrix 2

		Space Relations Test
Abstract Reasoning	Pearson correlation	0.845*
	Sig. (2-tailed)	0.000
		Gibson's Spiral Maze (error)
Gibson's Spiral Maze (time)	Pearson correlation	−0.515*
	Sig. (2-tailed)	0.005

* Correlation is significant at the 0.01 level (2-tailed).

A logical dispersion of groups is shown, plotting the individuals and the regression lines of the groups against the background of the Abstract Reasoning *versus* Space Relation plot (Table 3, plot 2). In other words, individuals of group 1 do not score below 25 on either one of the aptitude tests.

DISCUSSION

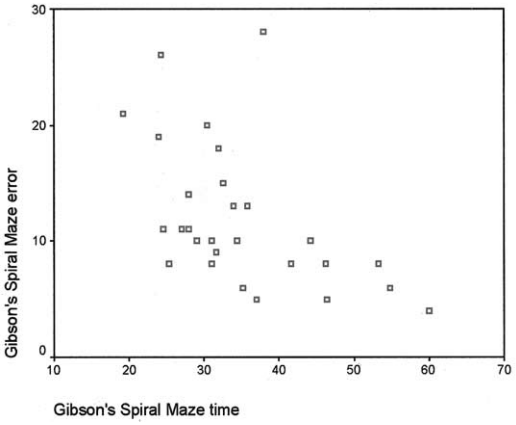
Recent interest in the selection of surgical trainees has been directed toward the use of aptitude tests [15]. Laparoscopic surgery, assuming a high degree of psychomotor abilities, could benefit from specifically oriented ability tests to aid in predicting career appropriateness for an aspirant laparoscopic surgeon.

It is important to realize what is meant by the term abilities itself. Abilities are innate, stable, and enduring aptitudes that an individual brings to the performance of tasks. Multiple perceptual motor abilities are identified, such as spatial orientation, information management, and manual dexterity. All individuals possess such abilities, but individuals have different patterns of strengths for their abilities, as defined by

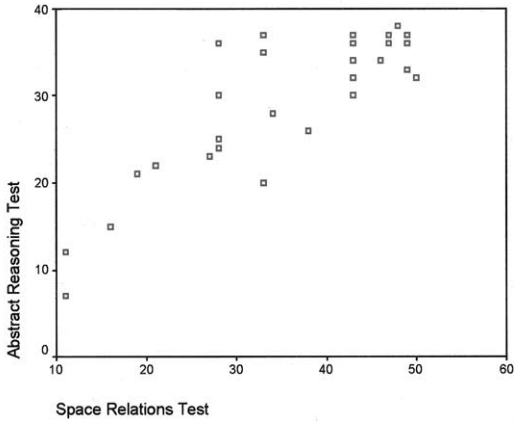
TABLE 3

Scatterplots

Plot 1: Gibson Spiral Maze



Plot 2: Space relation test versus Abstract Reasoning test



their genetic make-up. Skills describe an individual's proficiency at a particular task that has developed with, and is modified by, training and practice. An individual may acquire many skills through life, but each particular skill is based on a specific combination of the relatively small number of fundamental abilities. Thus, innate abilities underlie and are the limiting factor of the individual's performance of a certain task [16].

Different psychometric tests have been constructed for the identification of the level of an individual's particular ability. It must be stressed that the psychometric tests in our test battery were included because of their alleged capacity to assess a specific type of psychometric or psychomotor ability. Inevitably, test outcomes are related since the abilities they test are not mutually exclusive. There are no clear-cut groupings of psychometric or psychomotor abilities. Identification of psychometric and psychomotor abilities that underpin optimal endoscopic surgical performance would be most helpful for the selection of surgical trainees. Our study battery included four aptitude tests.

The Crawford Small Parts Dexterity Tester was selected because it has been reported to improve selection accuracy when used with other measures of selection for dental school students [17]. In the present study, this test seemed not to be significantly linked to Xitact test performance. It does seem to be significantly linked, however, to the Gibson Spiral Maze error test outcome. Although both tests are not able to explain Xitact test performance outcome, they cannot be regarded as useless for surgical assessment as they may discriminate for conventional psychomotor skill, not specifically addressed in this study.

The Gibson Spiral Maze test was chosen because of

TABLE 4

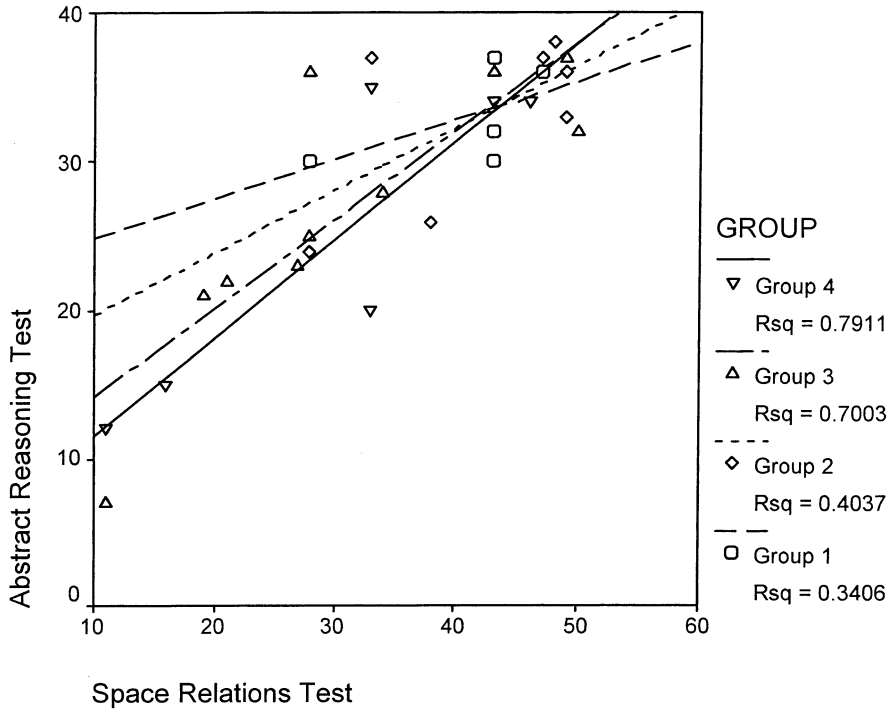
Aptitude Test Battery: Discriminative Power

Aptitude tests	Groups clustered	*N*	Mean rank	Mann–Whitney *U*	Asymp. Sig. (2-tailed)
Crawford Small Parts Dexterity Tester	Group 1 + group 2	12	12.33	70.000	0.329
	Group 3 + group 4	15	15.33		
Gibson Spiral Maze (time)	Group 1 + group 2	12	13.04	78.500	0.423
	Group 3 + group 4	16	15.59		
Gibson Spiral Maze (error)	Group 1 + group 2	12	14.46	95.500	0.982
	Group 3 + group 4	16	14.53		
Abstract Reasoning Test	Group 1 + group 2	12	18.42	49.000	0.029*
	Group 3 + group 4	16	11.56		
Space Relations Test	Group 1 + group 2	12	18.25	51.000	0.037*
	Group 3 + group 4	16	11.69		
	Total	28			

* Correlation is significant at the 0.05 level (2-tailed).

TABLE 5

Distribution Plot Examining Concurrent Validity



its reported contribution in assessing hand–eye coordination [18]. Earlier research showed faster executing time, comparing scores of surgical trainees with psychiatrist trainees. On the other hand, psychiatrist trainees had less error scoring. Thus, it has been difficult to assess and interpret the Spiral Maze's value in the testing for surgical aptitude. In our study, execution time and execution error are negatively correlated (the faster, the higher chance on errors). Since both parameters, time and error, are linked to the same concept, not much can be said studying only one parameter. Concepts are in itself not correlated to Xitact test outcome. Therefore, in this setting, authors feel the Gibson Spiral Maze test not to be of high value.

Visual–spatial ability is thought to be an important predictor of competence in specific surgical procedures. The Space Relations test, assessing visuospatial ability, proved to be a strong predictor for students resigning or delaying graduation for dentistry in a study of 1392 dental students [5]. Recent research in learning spatially complex surgical skill showed that visual–spatial ability is related to competence level and outcome after complex surgery. Wanzel and others conclude that visual–spatial ability testing can be used in the selection of surgical residents [19]. In our study, the Space Relations test is highly correlated to the outcomes of the Abstract Reasoning test. The plot combining the results of the Abstract Reasoning test with the Space Relations test outcomes seems to follow accurately the distribution of the novices' performance of Xitact. This is indicative of concurrent validity with Xitact surgical performance test scores.

The current study thus shows that two of the four tests included in the Aptitude Test Battery, i.e., the Abstract Reasoning Test and the Space Relation Test, have a predictive and selective value in identifying individuals who will achieve good laparoscopic surgical performance on the Xitact simulator. The Abstract Reasoning test is the only aptitude test which correlates directly with Xitact performance outcome. The test itself is highly correlated to the Space Relations test. There is a fan-like distribution of regression lines on the bivariate scatter plot of both aptitude tests. No outliers of performance are present among members of group 1. All members of this group score above 75% in correct answers on the abstract reasoning test. In contrast, members of group 4 are less able to perform both tests with good results. As a group, individuals in group 2 and 3 do indeed show intermediate test results. Thus, it is unlikely that persons with high innate psychomotor abilities are performing poorly on a combination of both the abstract reasoning test and the space relations test.

It must be said, however, that the study participants, all surgical novices, were pooled from a heterogeneous group of hospital interns (possibly interested

in a surgical specialty) and nonsurgical hospital residents (urology, emergency medicine, radiology, pulmonology, cardiology, anesthesiology, and internal medicine). It may be so that these participants will have lower scores on the Xitact simulator than, in fact, true novice surgical trainees. It is possible that self-selection, based on a person's own perception of his or her surgical abilities, leads to interest and a career in another medical specialty. It may also be a non-ability-related choice and merely a matter of (lack of) surgical career interest. In fact, surgeons themselves seem to be quite incapable of estimating their own performance in surgical simulation [20].

Nevertheless, performance scores on Xitact were objectively assessed. For the studied individual, comparing his or her Xitact score with his or her psychomotor aptitude test battery outcome is therefore valid, and in our study, a logic dispersion among performance groups is displayed. Extern validity, in terms of extrapolation of our results toward a *group* of aspiring surgical trainees, can be biased by rather an underestimate than an overestimate of innate abilities—as evaluated by Xitact outcome. Our test results must, therefore, be considered to be a prudent estimate of aspiring laparoscopic surgical trainee outcome.

Bearing the above in mind, our study results lead to the following preliminary step-wise selection recommendations, considering economic use of our study resources (e.g., the Aptitude Test Battery and Xitact simulator). Aspirant and novice laparoscopic surgeons can be tested using the Abstract Reasoning test and the Space Relations test. The group of participants scoring below 25 on either test, that is, an expected 36% of the population, would have to be monitored on Xitact to assess surgical task performance. The group of participants scoring above 35 on the Abstract Reasoning test and above 45 on the Space Relations test, that is, an expected 18% of the population, is unlikely to mal-perform on Xitact. The other 46% could very well benefit from Xitact simulation and assessment when the opportunity is present.

CONCLUSION

The use of aptitude testing for aspirant surgical trainees is not without debate. It is important to realize that a "good" surgeon is not a merely a product of a persons' knowledge and psychomotor abilities. Personality traits, such as interest, endurance, empathy, stress-resistance, and decision-making abilities are important and equally necessary to make up an all-round good surgeon. It is in the combination of these skills and personality variables the "ideal" surgeon must be sought.

Nevertheless, it is hard for an aspirant—laparoscopic—surgeon to compensate for a deficiency in specific psychomotor abilities. Previous studies show that surgeons are not good in estimating their own performance [20] and authors feel someone should at least have the opportunity to objectively assess these important features before deciding on a surgical laparoscopic apprenticeship. Such an opportunity could be offered by aptitude testing.

Our study shows that it is in the combination of a subset of aptitude tests that a surplus value can be seen. Therefore, it is surprising that studies focusing on aptitude test battery outcome have not regularly studied their test interaction and/or (cor-) relation.

Aspirant surgical trainees should be offered the possibility to undergo aptitude testing before embarking on a surgical traineeship. The advantages for both trainee and educator are evident. Information resulting from such tests can help the trainee to make an appropriate career decision, preparing him or her for future skills training or, in fact, revising initial aspirations. In any case, it is best to acknowledge aptitude problems early on in the selection, so that a trainee with a suboptimal set of innate abilities for a laparoscopic surgical career has basically two options. He or she can either choose to pursue a career in laparoscopic surgery nevertheless, knowing there will be heavier emphasis on skills training compared to his or her peers to attempt to achieve the same surgical end-quality. Or, it may be a better suitable option for that person to opt for one of the alternative careers in surgery or other area of medicine. For the educator, aptitude testing helps in tailoring specific skills training programs, offering the right amount of training to the individual.

The above is only possible once aptitude testing is considered to be valid and reliable.

The emerging era of medical virtual-reality simulation offers excellent opportunities for repetitive, reliable, and objective assessment and can thus double as a training, as well as a selection, instrument once properly validated. The Xitact LS500 laparoscopy virtual-reality simulator under study significantly, and in expected orderly fashion, correlates to the Abstract Reasoning Test of our Aptitude Test Battery. Our study established the concurrent validity of the Xitact LS500 with the combination of the Space Relations and Abstract Reasoning test measuring individual's visuospatial abilities.

REFERENCES

1. Muller, M. H. How does aviation find the ideal pilot? Suitability testing: Applicability to surgery? Methods for determining basic occupational suitability. *Zentralbl. Chir.* **124**(10): 889, 1999.
2. Carretta, T. R., and Ree, M. J. U.S Air Force pilot selection tests: What is measured and what is predictive? *Aviat. Space Environ. Med.* **67**(3): 279, 1996.
3. Dashfield, A. K., and Smith, J. E. Correlating fibreoptic nasotracheal endoscopy performance and psychomotor aptitude. *Br. J. Anaesth.* **81:** 687, 1998.

4. Price, M. Selection of medical students. *Br. Med. J.* **324:** 1170, 2002.
5. Smith, B. G. A longitudinal study of the value of a spatial relations test in selecting dental students. *Br. Dent. J.* **167**(9): 305, 1989.
6. Schueneman, A. L., *et al.* Neuropsychologic predictors of operative skill among general surgery residents. *Surgery* **96**(2): 288, 1984.
7. Deary, I. J., Graham, K. S., and Maran, A. S. Relationships between surgical ability ratings and spatial abilities and personality. *J. R. Coll. Surg. Edinb.* **37**(2): 74, 1992.
8. Wanzel, K. R., Ward, M., and Reznick, R. K. Teaching the surgical craft: From selection to certification. *Curr. Probl. Surg.* **39**(6): 574, 2002.
9. Risucci, D. A. Visual spatial perception and surgical competence. *Am. J. Surg.* **184**(3): 291, 2002.
10. Simon, J. F., and Chambers, D. W. The search for a profile of aptitudes that characterize dentists. *J. Dent. Educ.* **56**(5): 317, 1992.
11. Francis, N. K., *et al.* Performance of master surgeons on standard aptitude testing. *Am. J. Surg.* **182**(1): 30, 2001.
12. Schijven, M. P., Jakimowicz, J. J. The learning curve on the Xitact LS500® laparoscopy simulator: Profiles of performance. *Surg. Endosc.* **18**(1): 121, 2004.
13. Schijven, M., and Jakimowicz, J. Construct validity: Experts and residents performing on the Xitact LS500 laparoscopy simulator. *Surg. Endosc.* **17:** 803, 2003.
14. Schijven, M., and Jakimowicz, J. Face-, expert- and referent validity of the Xitact® LS500 Laparoscopy Simulator. *Surg. Endosc.* **16:** 1764, 2002.
15. Graham, K. S., and Daery, I. J. A role for aptitude testing in surgery? *J. R. Coll. Surg. Edinb.* **36**(2): 70, 1991.
16. Dashfield, A. K., *et al.* Correlation between psychometric test scores and learning tying of surgical reef knots. *Ann. R. Coll. Surg. Engl.* **83:** 139, 2001.
17. Boyle, A. M., and Santinelli, J. C. Assessing psychomotor skills: The role of the Crawford Small Parts Dexterity Test as a screening instrument. *J. Dent. Educ.* **50**(3): 176, 1986.
18. Harris, C. J., Herbert, M., and Steele, R. J. Psychomotor skills of surgical trainees compared with those of different medical specialties. *Br. J. Surg.* **81**(3): 382, 1994.
19. Wanzel, K. R., *et al.* Effect of visual-spatial ability on learning of spatially-complex surgical skills. *Lancet* **359**(9302): 230, 2002.
20. Schijven, M. P., Jakimowicz, J., and Schot, C. The Advanced Dundee Endoscopic Psychomotor Tester (ADEPT) objectifying subjective psychomotor test performance. *Surg. Endosc.* **16:** 943, 2002.