World J. Surg. 25, 1478–1483, 2001
DOI: 10.1007/s00268-001-0133-1

WORLD
Journal of
SURGERY

Objective Psychomotor Skills Assessment of Experienced, Junior, and Novice Laparoscopists with Virtual Reality

Anthony G. Gallagher, Ph.D., Karen Richie, B.Sc., Neil McClure, M.D., Jim McGuigan

The Northern Ireland Centre for Endoscopic Training and Research and The Queen's University of Belfast, Malone Road, Belfast BT7 1NN, United Kingdom

Published Online: September 28, 2001

Abstract. Objective assessment of psychomotor skills in surgery is now a priority; however, this assessment is difficult to achieve because of measurement difficulties associated with the reliability and validity of assessing surgical skills in vivo and in the laboratory. In this study virtual reality (VR) was used to overcome these problems in the objective psychomotor assessment of senior, junior, and novice laparoscopists. Twelve experienced laparoscopic surgeons (performed >50 Minimal Access Surgery (MAS) procedures), 12 inexperienced laparoscopic surgeons (<10 MAS procedures), and 12 laparoscopic novices (no MAS procedures) participated in the study. Each subject completed all six tasks of the Minimally Invasive Surgical Trainer; Virtual Reality (MIST VR). In comparison to the other groups, experienced laparoscopic surgeons performed the tasks significantly ($p < 0.01$) faster, had a lower error rate, were more economic in their movement of surgical instruments and in the use of diathermy. As a group they also showed greater consistency in their performance. MIST VR distinguished between the three groups of laparoscopists. VR provides a useful objective assessment tool for evaluating psychomotor skills for laparoscopic surgery.

Traditionally, surgeons have been expected to adapt through real life experience to the various environmental and equipment constraints of the operating theater. With the advent of Minimal Access Surgery (MAS) or laparoscopic surgery, more explicit attention has been paid to "human factors" in surgery, particularly to the ergonomic shortcomings of the operating environment and the limits of individuals' psychomotor, cognitive, and perceptual faculties [1]. The advantages of the laparoscopic method for the patient have been well documented [2]. But for the surgeon, laparoscopic surgery is more demanding, requiring greater concentration than open surgery. Operating times are longer, with increased surgical fatigue and stress because of the remote intervention associated with the laparoscopic approach. "Surgical fatigue syndrome" is characterized by mental exhaustion, increased irritability, impaired surgical judgment, and reduced dexterity [3]. Furthermore, research has indicated that laparoscopic surgery has been associated with a higher rate of complications in comparison to open surgery [4, 5].

Correspondence to: A.G. Gallagher, Ph.D., School of Psychology, The Queen's University of Belfast, Malone Road, Belfast BT7 1NN, UK, e-mail: ag.gallagher@ qub.ac.uk

Well-known perceptual and cognitive problems contribute to these patterns. Tactile feedback during surgery is considerably degraded. A second purely visual problem associated with the learning and practice of laparoscopic surgery is that the laparoscopic surgeon needs to form visual impressions of a three-dimensional structure—consisting of organs and instruments—from a two-dimensional television monitor. That presents a problem which is often described as loss of binocularity [6], but 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 they (and other cues related to lighting and texture) yield a conclusion that is inimical to surgery: they specify that the structures in view form a single surface, virtually flat and usually vertical. A surgeon has to set aside that conclusion; to register the information carried by subtler "pictorial" cues; and to reconstruct the structure that they specify despite the incompleteness of the information that they provide.

A third and more fundamental problem for laparoscopic surgeons is that perceptual–motor correlation is inverted: that is, when the surgeon's hand moves to the patient's right, the working end of the instrument the surgeon is holding moves to the left of the screen. The inversion is due to the "fulcrum effect" of the patient's abdominal wall. This effect is known to have a major detrimental effect on the performance of trainee laparoscopic surgeons [7]. In another study we have demonstrated that with training and practice surgeons will automate to the fulcrum effect [8]. Individuals have traditionally learned this hand-eye coordination on bench models or animals. Dath and Reznick [9] have suggested that practical as these have been they are limited in the quality and quantification of the surgeons' performance. They propose an objective structures assessment of technical skill (OSATS) examination format. Virtual reality simulators may offer another alternative.

Computer-generated simulators have been used for many years in the training and assessment of pilots and boat crews. The last 10 years have seen numerous attempts to introduce virtual reality imaging into clinical medicine. The term virtual reality refers to "a computer generated representation of an environment that allows sensory interaction, thus giving the impression of actually being there" [10].

In 1991, Satava and Lanier created the first virtual reality surgical simulator [11]. It was designed primarily as a training tool and focused on the abdominal area only. However, because of the limitations in computer processing capacity and resolution of liquid crystal displays, the virtual abdomen was relatively simple in appearance. As the VR industry has matured, simulators have become more sophisticated and complex with some generating near-photorealistic images and haptic feedback [12].

One of the most recent developments in laparoscopic surgery simulation is the Minimally Invasive Surgical Trainer Virtual Reality (MIST VR). The virtual reality system was designed to develop and assess minimally invasive surgical skills using advanced computer technology, which could be easily operated by both tutor and trainee. The system comprises a frame holding two standard laparoscopic instruments, which are linked to a Pentium PC. The MIST VR constructs a virtual environment on the screen, which shows the position and movement of the instruments in real time. An accurately scaled operating volume of 10 cm^3 is represented by a three-dimensional cube on the computer screen. The overall image size and the sizes of the target object can be varied for different skill levels. Targets appear randomly within the operating volume according to the task and can be "grasped" and "manipulated" [13].

In training mode, the program guides the trainee through a series of six tasks which progressively become more complex, enabling the development of the hand-eye motor coordination essential for the safe clinical practice of laparoscopic surgery. Each task is based on an essential surgical technique employed in MIS. The completion of the tasks requires the use of both hands, which is an important skill for laparoscopic surgery.

In three other studies we have demonstrated the value of MIST VR as a training device. The results of these studies showed that subjects who received training on MIST VR significantly outperformed case-matched control groups [14–16].

Although there is a growing body of evidence on the value of MIST VR as a training device there has been relatively little work on it as an assessment device. Objective assessment of psychomotor ability and learnability for laparoscopic surgery has made significant advances in the last few years [17–19]. It is also a matter of some importance to professional organizations such as the Royal College of Surgeons [20]. If MIST VR is a valid laparoscopic assessment device it should be capable of distinguishing between the objectively scored performance of experienced laparoscopic surgeons, junior laparoscopic surgeons, and laparoscopic novices. The purpose of the study reported here was to establish whether experienced laparoscopic surgeons perform better than junior laparoscopic surgeons and laparoscopic novices.

Methods

Participants

Twelve experienced laparoscopic surgeons who had performed >50 laparoscopic operations participated (mean age = 39, range 30–52), 12 less experienced surgeons who performed >1 but <10 laparoscopic operations (mean age = 29, range 25–32), and 12 individuals with no laparoscopic operative experience (mean age = 21, range 19–22) participated.

Apparatus

The MIST VR system used in this study was based on 2000 MHz Pentium PC running Windows 95 with 32 Mb RAM and 1.6Gb hard disk, a Martrox Mystique 4MB video card. The laparoscopic interface was a standard Immersion Corporation (San Jose, CA, USA) unit, with the addition of a foot pedal for the diathermy tasks. This jig, containing two laparoscopic instruments held in position-sensing gimbals provided 5 degrees of freedom. The trials ran MIST VR Version 1.2, which utilized the WorldToolKit Version and Microsoft Direct 3D Version 3 graphics libraries. Frame rates averaged around 15 frames per second (fps). This provides real-time translation of the instrument movements to the graphic display on a 17′ color monitor. A 3-D cube on the computer screen represents an accurately scaled operating volume of 10 cm^3. The image zoom and size of target objects can be varied. Targets appear randomly within the operating volume according to the skill task and can be "grasped" and "manipulated." Each of the different tasks is recorded exactly as it is performed and could, therefore, be accurately and reliably assessed. For the studies reported here the monitor was positioned at eye level with the operating jig at standard surgical height between the participant and the monitor.

Procedure

All subjects received supervised MIST VR testing and completed all six levels of the simulation tasks. Testing was completed in a quiet room near the operating theaters. There were five measures of participants' performance. These were the number of seconds required to complete all six MIST VR tasks, number of errors, economy of movement of the left instrument, economy of movement of the right instrument, and economy of diathermy use during tasks five and six. All of the tasks were completed five times per trial for each hand.

Measurement Definition

Time. The time subjects spent on the different tasks was measured from when they started on the task until the time that they completed the last sequence of movements. The clock stopped automatically at the end of the task. The time between the different tasks (i.e., changing from task One to Two, etc.) was not measured.

Economy of Movement of Left and Right Instruments. Economy of movement was assessed for the left and right instrument's as the proportion of the distance travelled by the left (or right) instrument tip (or working end) as it exceeded its optimal distance. This measures how far past the virtual object the virtual instrument went when completing the task.

Error. Average error was measured as the number of errors per task segment.

Economy of Diathermy Use (Tasks 5 and 6 Only). The total burn time was used as a measure of economy of diathermy score for both tasks 5 and 6. These were the only two tasks to employ diathermy as a task component.

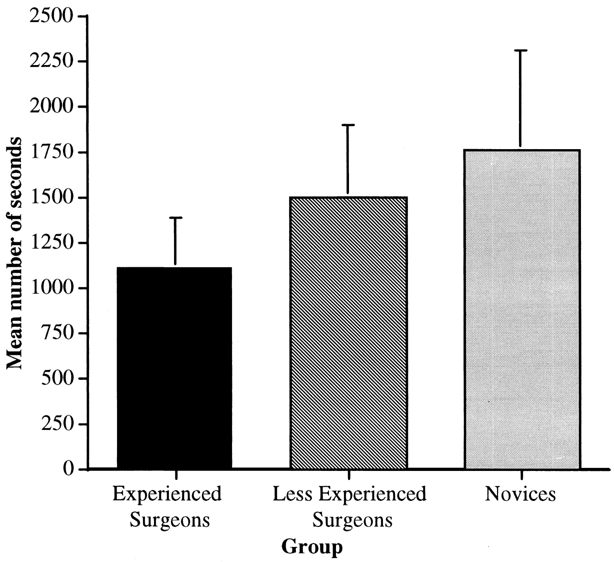

Fig. 1. Mean and standard deviation number of seconds taken by the three groups to complete the MIST VR tasks.

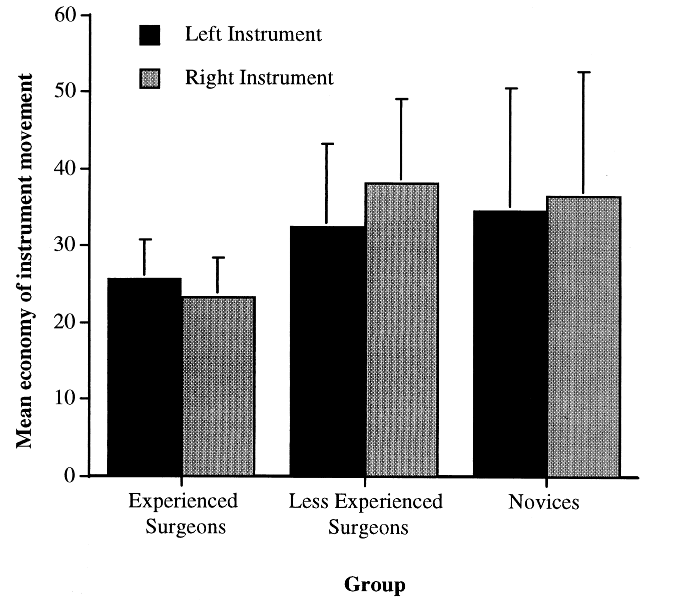

Fig. 2. Mean and standard deviation for the economy of movement of the left and right instruments by the three groups.

Results

For ease of analysis participants' responses were summed across the six tasks for the time, error, and economy of movement measures. Economy of diathermy was summed for the two tasks in which it was a feature of that is, tasks 5 and 6.

Figure 1 shows the mean number of seconds to complete the MIST VR tasks by the two groups of laparoscopic surgeons and the laparoscopic novices. Overall, there was a significant difference between the groups performance ($F(2,33) = 7.21$, $p = 0.0025$). The experienced surgeons performed the tasks significantly faster than the less experienced surgeons ($p < 0.05$) and the laparoscopic novices ($p < 0.001$). Although the novices took longer to complete the MIST VR tasks than the less experienced surgeons, this difference was not found to be statistically significant. The less experienced and laparoscopic novice groups also showed more variability in their performance as indicated by their standard deviation scores.

The economy of movement measures of the left and right instruments are shown in Figure 2. Although there were some differences among the groups in terms of which hand they performed best with these were not statistically significant. There was a statistically significant difference between the groups in economy of movement performance of the left ($F(2,33) = 6.04, p = 0.0058$) and right instruments ($F(2,33) = 5.47$, $p = 0.009$). The experienced surgeons were significantly more economic in the movement of the left instrument than the less experienced surgeons and the novices ($p < 0.01$). Although the less experienced surgeons performed slightly better than the laparoscopic novices, this difference was not significant. The same pattern was true for the right instrument (experienced surgeons versus less experienced surgeons, $p < 0.05$, and experienced surgeons versus novices, $p < 0.01$). Experienced surgeons were also more consistent in their economy of movement as indicated by the smallest standard deviations of the three groups.

The mean score for the MIST VR error measure showed that the experienced surgeons had a lower error score (mean = 18.12, S.D. = 5.4) than the less experienced surgeons (mean = 29.95, S.D. = 14.4) and the novices (mean = 26.75, S.D. = 21.2). However, one of the most notable observations from these data was the variability of scores for the three groups. This is shown in the box plot in Figure 3, which shows the 10th, 25th, 50th, 75th, and 90th percentiles for each group. Experienced surgeons' performance showed considerable heterogeneity in comparison to the other two groups. Differences among the groups were examined for significance with ANOVA where equality of variance was *not* assumed. The results showed that the experienced surgeons had a significantly lower error rate than the less experienced surgeons ($p < 0.05$). Other differences were not statistically significant.

The economy of diathermy data showed a similar pattern to the

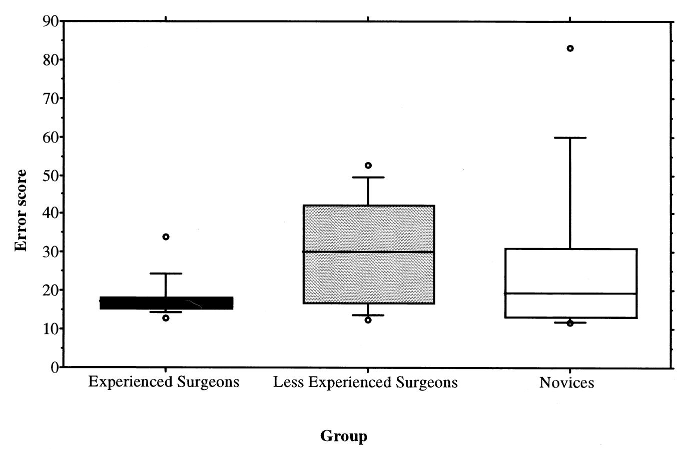

Fig. 3. Median, 10th, 25th, 75th, and 90th percentile error scores for the three groups.

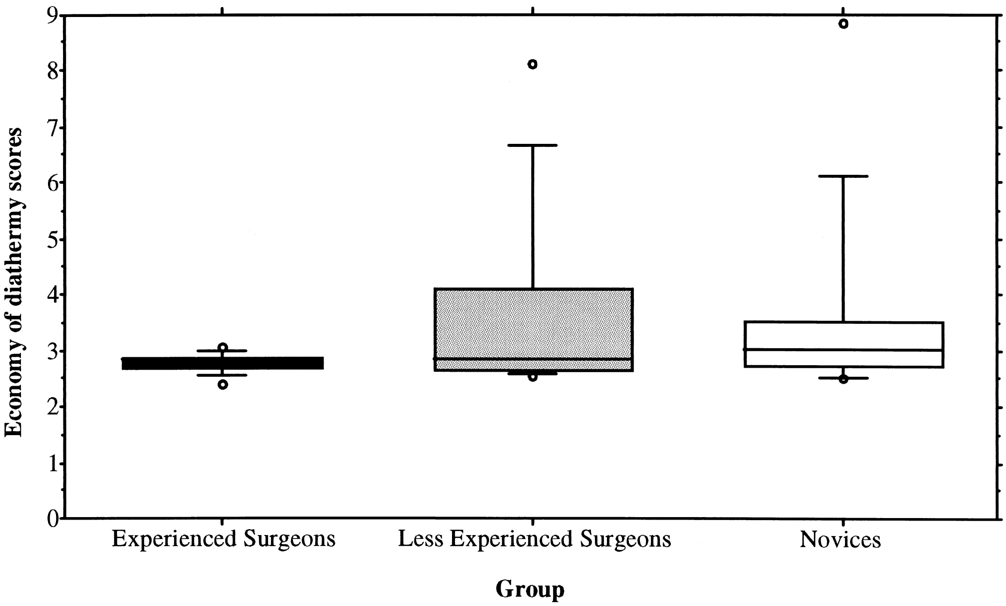

Fig. 4. Median, 10th, 25th, 75th, and 90th percentile error scores for the three groups.

error measure and is shown in the boxplot in Figure 4. There were no significant differences among the mean performance of the groups. Again, the most notable observation about these data is the low variability in the experienced surgeons' scores in comparison to the other two groups. The standard deviations for the three groups were; experienced surgeons, S.D. = 0.18; less experienced surgeons, S.D. = 1.72; laparoscopic novices, S. D. = 1.8. Using the experienced surgeons' standard deviation score as the comparison metric unit, the less experienced surgeons' and laparoscopic novices' standard deviation scores are 960% and 1000% greater, respectively.

Discussion

The study reported here set out to investigate whether it was possible to distinguish between laparoscopic surgeons of different levels of experience on the basis of objective measurements generated from their performance on a computer from a series of "virtual" tasks. A fundamental psychomotor problem for trainee laparoscopic surgeons is the difficulty learning or acquiring the hand-eye coordination to become a laparoscopic or endoscopic surgeon. One of the major problems in the acquisition of these skills is the "fulcrum effect" of the body wall on instrument manipulation [7]. The trainee paying closer attention to the task being performed cannot correct for this problem. Due to the cognitive demands of endoscopic tasks on trainees it is almost impossible simultaneously to pay attention both to the task and to hand coordination because of the attentional load of dual task performance. Most trainee endoscopists will eventually automate to the fulcrum effect [8]. This means that they will be able to coordinate their hands without consciously having to attend to their coordination. However, the rate of automation will probably vary among individuals.

In this study the psychomotor performance of experienced surgeons was faster, they made fewer errors, and were more economic in the movement of instruments and in their use of diathermy. Experienced surgeons were also more consistent in their performance as indicated by smaller standard deviation scores. Across all the measures the standard deviations of the experienced surgeons was small. This was particularly marked in the error measure and the diathermy scores.

Measures of variance are important performance parameters to look at when investigating skill acquisition [21]. The goal of training is to increase performance levels but also to decrease variability in performance. This was demonstrated most clearly with the error and diathermy scores. Perhaps the heterogeneity observed in the performance of experienced surgeons is a metric of what surgeons refer to as a "safe pair of hands", that is, economy and grace of movement.

The diathermy measure was the only one to show no significant differences among the groups; however, there were clearly different score patterns among the groups. Diathermy, when utilized during in vivo surgery, can be used for coagulation of a hemorrhage or resection of tissues. The important point about the use of diathermy is that it uses an electric charge at the end of a surgical instrument. Diathermy is extremely effective at coagulation or cutting tissue. However, the use of diathermy must be conservative because this current is being used in a moist enclosed environmental space (e.g., the abdomen). It is possible that the electrical charge could "arc" and damage other tissue that the surgeon was not aiming at, for example bowel or liver. This situation can be extremely dangerous because the surgeon is operating with minimal access and may not see or be aware of the damage. Experienced surgeons are very aware of this. The data from the study reported here support this conclusion. Experienced surgeons were conservative in their use of diathermy. In contrast, the junior surgeons and laparoscopic novices were less conservative as demonstrated by their score variability.

Assessment of surgeons will become of increasing importance in the next 5 years post "Bristol Case." In Bristol (UK) two pediatric cardiac surgeons were very publicly removed from the medical register for complications that were "unacceptably high."

The president of the Royal College of Surgeons (Mr. Barry Jackson) has suggested that the assessment and training of surgeons should be revisited [20]. Many surgeons believe that crucial aspects of surgical ability such as psychomotor ability or spatial skills cannot be assessed and that performance on simulators, be they virtual reality or other, simply indicates how that individual performs on the simulator and no more. They believe that the only way to assess surgical competence or whether a surgeons has a "safe or good pair of hands" is by observation by an experienced surgeon. However, others such as Darzi, Smith, and Taffinder [17] believe that the need to find reliable methods of measuring surgical competence is now an urgent matter.

In many respects laparoscopic surgeons are receptive to the idea of objective assessment, probably because of the difficulties that were associated with the introduction of laparoscopic surgery, that is, increased morbidity and mortality [22]. They are also aware that laparoscopic surgery requires very different skills from those of general surgery and that some surgeons may not have or be able to learn those abilities. This has been reported to us in a number of personal communications from experienced and scientifically sophisticated surgeons from the United Kingdom and the United States of America. These surgeons suspect that approximately 5%, 10% of junior surgeons simply do not have the requisite psychomotor, perceptual, or visiospatial ability to become a safe or competent laparoscopic surgeon. However, this claim has never been objectively tested. Furthermore, some benchmark measure would need to be established against which to assess individuals, preferably on an international basis and with much larger numbers than were used in this study. Another important point is that any battery of tests that would purport to assess skills that are important to laparoscopic surgery would also need to assess objectively the ability of the individual to reconstruct three-dimensional images from two-dimensional perceptual cues and their ability to deal with these spatial images. However, metric devices such as MIST VR may facilitate the objective assessment of gross and subtle aspects of psychomotor performance. Ultimately these are empirical questions. They are currently being aggressively pursued in our laboratory and in laboratories in the United Kingdom and United States of America.

Conclusions

In this study it has been shown that it is possible to distinguish between the objectively scored performance of experienced laparoscopic surgeons and less experienced surgeons and controls on virtual reality tasks. These differences were apparent in both mean performance, economy of instrument usage, and score variability. These are important metrics in the objective assessment of skill acquisition that could be applied internationally with standardized virtual tasks.

Résumé

Buts: L'évaluation objective des gestes psychomoteurs en chirurgie est maintenant une priorité, cependant, celle-ci est difficile à réaliser en raison de la fiabilité et de la validité de l'évaluation de la chirurgie in vivo ainsi que dans le laboratoire. Dans cette étude, on a fait appel à la réalité virtuelle (RV) pour essayer de résoudre ces problèmes dans l'évaluation psychomotrice objective des laparoscopistes novices, senior et juniors. Méthodes: Douze chirurgiens laparoscopiques expérimenté (>50 procédés de chirurgie mini-invasive (CMI)), 12 chirurgiens laparoscopiques non-expérimentés (<10 procédés CMI) et, 12 novices en laparoscopic (aucun procédé CMI) ont participé à cette étude. Chaque sujet a accompli six gestes avec un traneur Minimally Invasive Surgical Trainer Virtual Reality (MIST VR). Résultats: Par comparaison aux autres groupes, les chirurgiens expérimentés ont accompli leurs gestes de façon significativement plus vite ($p < 0.01$), avec moins d'erreurs; ils étaient plus économiques dans leurs mouvements des instruments chirurgicaux et dans l'utilisation de la coagulation. Comme un groupe, ils étaient plus consistants dans leur performances. Conclusion: Le système MIST VR a pu distinguer entre les trois groupes de chirurgiens laparoscopistes. Le MIST VR fournit un outil d'évaluation objective pour évaluer les gestes psychomoteurs en chirurgie laparoscopique.

Resumen

Objetivos: En la actualidad un objetivo prioritario en cirugia es la valoración objetiva de la destreza psicomotora. Tanto en cirugia "in vivo" como en la experimental, la valoración fidedigna de ésta es dificil. En este estudio se utilizó la realidad virtual (V. R.) para, obviando estas dificultades, evaluar de forma objetiva las habilidades psicomotoras de laparoscopistas avanzados, con menor experiencia y sin experiencia alguna. Métodos: 12 cirujanos laparoscopistas experimentados (>50 intervenciones MAS), 12 inexpertos cirujanos laparoscopistas (<10 intervenciones MAS) y 12 laparoscopistas novicios (ninguna intervención MAS) participaron en este estudio. Cada uno de ellos efectuó 6 ejercicios de cirugia minimamente invasiva utilizando la realidad virtual (MIST VR). Resultados: Los laparoscopistas con experiencia realizaron dichos ejercicios más rápidamente que los de los otros dos grupos ($p<0.01$), el porcentaje de errores fue bajo y fueron más parcos en el manejo de los instrumentos quirúrgicos y en el empleo de la diatermia. Como grupo, demostraron gran coherencia en su actuación. La VR, en cirugía laparoscópica, es un instrumento muy útil para evaluar con exactitud las aptitudes y destreza psicomotoras.

References

1. Berguer, R.: Ergonomics in the operating theatre. Am. J. Surg. *171*: 385, 1996
2. Prasad, A., Foley, R.J.E.: Day case laparoscopic cholecystectomy: a safe and cost effective procedure. Eur. J. Surg. *162*:4, 1996
3. Cuschieri, A.: Whither Minimal Access Surgery: tribulations and expectations. Am. J. Surg. *169*:9, 1995
4. Deziel, D., Milikan, K. W., Economou, S. G., Doolas, A., Ko, S. T., Aairan, M. C.: Complications of laparoscopic cholecystectomy: a national survey of 4, 292 hospitals and an analysis of 77, 604 cases. Am. J. Surg. *165*:9, 1993
5. Peters, J. H., Ellison, E. C., Innes, J. T.: Safety and efficacy of laparoscopic cholecystectomy. Ann. Surg. *312*:3, 1991
6. Reinhardt-Rutland, A. H., Gallagher, A. G.: Visual depth perception in minimally invasive surgery. In Contemp Ergon S. A. Robertson editors, London: Taylor Francis, 1996, pp. 531–536
7. Gallagher, A. G., McClure, N., McGuigan, J., Ritchie, K., Sheehy, N.: An ergonomic analysis of the 'Fulcrum Effect' in endoscopic skill acquisition. Endoscopy *30*:617, 2001
8. Crothers, I., Gallagher, A.G., McClure, N., James, D.T.D., McGuigan, J.: Experienced surgeons are automated to the fulcrum effect: an ergonomic demonstration. Endoscopy *31*:365, 1999

9. Dath, D., Reznick, R.: Objective skills assessment: bench model solutions. In Surgical Competence: Challenges of Assessment in Training and Practice, Royal College of Surgeons of England, editors. London, RCS & Smith and Nephew, 1999, pp. 52–66
10. Coleman, J., Nduka, C.C., Darzi, A.: Virtual reality and laparoscopic surgery. B.J.S. *81*:1709, 1994
11. Satava, R.M.: Virtual reality surgical simulator; the first steps. Surg. Endosc. *7*:203, 1993
12. Satava, R.M., Jones, S.B.: Virtual reality. In Cybersurgery: Advanced Technologies for Surgical Practice, R. M. Satava editor, New York, Wiley-Liss, pp. 75–96, 1998
13. Wilson, M.S., Middlebrook, A., Sutton, C., Stone, R., McCloy, R.F.: MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance. Ann. R. Col. Surg. Engl. *79*:403, 1997
14. Gallagher, A.G., McClure, N., McGuigan, J., Crothers, I, Browning, J.: (1999) Virtual reality training in laparoscopic surgery: a preliminary assessment of Minimally Invasive Surgical Trainer Virtual Reality (MIST VR). Endoscopy *31*:310, 1999
15. Gallagher, A.G., Hughes, C., Reinhardt-Rutland, A.H., McGuigan, J., McClure, N. (2000) A case control comparison of traditional and VR training in laparoscopic psychomotor performance. Minim. Invasive Ther. All Techn. 9:347
16. Jordan, J.A., Gallagher, A.G., McGuigan, J., McGlade, K., McClure, N. (2000): A comparison between randomly alternating imaging, normal laparoscopic imaging and virtual reality training in laparoscopic psychomotor skill acquisition. Am. J. Surg. 180:208
17. Darzi, A., Smith, S., Taffinder, N.: Assessing operative skill. B. M. J. *318*:887, 1999
18. Rosser, J.C., Rosser, L.E., Savalgi, R.S.: Skill acquisition and assessment for laparoscopic surgery. Arch. Surg. *132*:200, 1997
19. Rosser, J.C., Rosser, L.E., Savalgi, R.S.: Objective evaluation of laparoscopic surgical skill programme for residents and senior surgeons. Arch. Surg. *133*:657, 1998
20. Jackson, B.: Foreword. Royal College of Surgeons of England. Surgical Competence: Challenges of Assessment in Training and Practice. London, RCS & Smith and Nephew, 1999, p. 5
21. Adams, J. A.: Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. Psychol. Bull. *101*:41, 1987
22. The Southern Surgeons Club: The learning curve for laparoscopic cholecystectomy. Am. J. Surg. *170*:55, 1995