Technique

Learning rate for laparoscopic surgical skills on MIST VR, a virtual reality simulator: quality of human-computer interface

Aafia Chaudhry*, Christopher Sutton[†], Jonathan Wood*, Robert Stone[‡], Rory McCloy*

* *North of England Wolfson Centre for Minimally Invasive Therapy, Manchester Royal Infirmary, Manchester, UK*
[†] *Virtual Presence Ltd, London, UK*
[‡] *Virtual Presence Ltd, Manchester, UK*

Acquiring laparoscopic surgical skills involves initial learning of cognitive and motor skills followed by refinement of those skills. The successful use of a virtual reality simulator depends on the quality of the interface for the human-computer interaction and this can be determined by the initial learning rate. MIST VR, a part-task virtual reality laparoscopic simulator, provides objective assessment of psychomotor skills and can generate an overall score for performance, based upon errors made and time taken for six different tasks. This study analysed the rate of early task/instrument/computer familiarisation on consecutive scores achieved by surgically experienced and naive individuals. Eleven surgeons, 18 medical students and seven non-medical personnel were tested on the simulator up to ten consecutive times, within a 2-week period. Performance data from every task and repetition were analysed to obtain individual scores of task performance. The calculation of overall score penalised errors far more heavily than total time taken, with high scores indicating poor performance. The surgeon-computer interface generated a rapid and significant early familiarisation curve up to the third session on the simulator, with significant reductions in both time taken and total contact errors made. These results suggest that MIST VR represents a high quality interface. Surgeons scored consistently and significantly better than other subjects on all tasks. For surgically naive individuals, it was possible to predict the level of laparoscopic skills performance that would be attained after overcoming initial simulator learning curve, by studying their initial score. Overall scores reflected surgical experience and suggest that the simulator is measuring surgically relevant parameters.

Correspondence to: Mr Rory McCloy, University Department of Surgery, Manchester Royal Infirmary, Oxford Road, Manchester M13 9WL, UK.   Tel: +44 161 276 4534; Fax: +44 161 273 3428; E-mail: R.McCloy@man.ac.uk

MIST VR provides a validated and much needed method for objective assessment of laparoscopic skills, for a variety of surgical disciplines.

Key words: Laparoscopy – Simulation – Virtual reality – Surgical training – Surgical performance – Minimal access surgery – Human-computer interaction

The past two decades have seen enormous developments in the field of minimally invasive therapy. For some time, laparoscopic cholecystectomy has succeeded open surgery as the preferred treatment for symptomatic gallstones, but, at present, there are no existing standards that must be met by surgeons in order to practise this technique.[1] The skills employed in successful laparoscopic surgery require different hand-eye co-ordination and skills than those developed in conventional open surgery. At present, trainees are still principally instructed by classical surgical apprenticeship that may become inappropriate within the next decade.[2]

The report of the working party on specialist medical training[3] has suggested that 'the doctor has satisfactorily completed specialist training based on an assessment of competence'. This is echoed by the European Association for Endoscopic Surgery that has defined guidelines relating to the training and practice of minimally invasive surgery and also calls for certification of competence in performing laparoscopic procedures.[4] However, to date, there has been no objective method of formally assessing a trainees performance or progress in acquiring surgical techniques, open or laparoscopic.

Given the variability in clinical skills of doctors entering surgical training programs, their desire to receive objective feedback relating to their skills[5] and also their progress during training, MIST VR, (minimally invasive surgical trainer – virtual reality), was developed by a joint venture between the North of England Wolfson Centre for Minimally Invasive Therapy, and Virtual Presence Ltd.[6,7]

Acquiring laparoscopic surgical skills involves initial learning of cognitive and motor skills followed by refinement of those skills. Due to the complex nature of human anatomy and technological limitations in simulating it, few attempts have been made to introduce computer simulation into surgical training.[8] MIST VR provides a new way of training and assessing laparoscopic surgical skills and surpasses the use of basic laparoscopic training boxes, in which inanimate objects are placed and manipulated using normal instruments. The use of such trainers fails to provide subjects with any clear indication of their level of

manual dexterity, their progress with training or even a comparison of performance with their peers.

MIST VR comprises two standard laparoscopic instruments linked via a jig to a PC. The instruments are held in position-sensing gimbals with 5 degrees of freedom and the movement of the instruments is displayed to the user via a propriety real-time graphical display package.[7] Six tasks have been developed for the trainer which were derived from an ergonomic assessment of manoeuvres performed during laparoscopic surgery (Table 1).[6] When a subject performs a task, every movement made by the instruments is recorded including any contact errors. A spreadsheet has been developed which provides scores for individual tasks and for right and left handed performance separately. Hence, a full record of that trainee's performance is logged and can be monitored over successive training sessions.

As yet, little is understood about the specific factors that influence a subject's performance, hence overall score on the simulator. This study sets out to: (i) analyse the human-computer interactions that occur during initial task/instrument/computer familiarisation when subjects are tested on MIST VR, on a repeated basis; and (ii) to determine, whether or not, there exists an early learning curve for the simulator. The quality of the interface, MIST VR, can be assessed by the rate of the early learning of the tasks involved.[9] If the initial learning curve is rapid, then this indicates a high quality human-computer interface. Furthermore, analyses of the participants' learning curves will determine how many examination sessions are necessary to evaluate an individual's cognitive and motor skills reliably. Testing of both surgically experienced and naïve subjects has been undertaken in order to validate the overall performance scores achieved on the simulator. This objective assessment of performance has significant implications for the future assessment of manual dexterity and hand-eye co-ordination in the context of laparoscopic skills, for both surgical trainers, trainees and the length of surgical training of those individuals contemplating a career in minimally invasive surgery across a variety of disciplines.

Subjects and Methods

Eighteen subjects were recruited into this study. Of these, 11 were surgeons (BST grade and above) and seven were non-surgeons (other hospital staff). In the surgically experienced group, there were nine male and two female subjects, (age range 25–35 years), and in the non-surgeons there were four male and three female subjects, (age range 22–41 years). All the subjects were right-handed.

All participants in the study had no previous contact with MIST VR. When they were approached to enter the study, a standardised, brief explanation regarding MIST VR was given and it was explained that an initial visit entailing approximately 40 min would be required and all subsequent visits would be around 20 min duration. Participants were required to attend a minimum of six or a maximum of ten visits, over the course of two working weeks.

At the initial visits, all participants were guided through the equipment and tasks using the MIST VR demonstrator protocol and were given the opportunity to practise the tasks with each hand. They then completed a questionnaire regarding general demography and another questionnaire that determined their level of right-handedness. Subjects then attempted all six tasks (with five repetitions for each hand), in 'exam

Table 1 Description of virtual tasks

Task no
1   Acquire randomly placed object   Accurate placement within a 3-D frame (one-handed task)

2   Acquire randomly placed object
    Transfer object to opposite instrument
    Accurate placement within a 3-D frame

3   Traverse a cylindrical object in steps using alternate   instruments

4   Acquire randomly placed object
    Withdraw instrument from operating field
    Accurately re-insert instrument to remake contact with
      object

5   Precise diathermy of 3 nodes on surface of randomly   positioned objects using floor pedal (one-handed task)

6   Acquire object
    Maintain position within 3-D frame
    Touch with other instrument
    Withdraw second instrument from operating field
    Re-insert as diathermy tool
    Diathermy 3 nodes on object surface whilst maintaining
      position within frame with other instrument

mode'. The six tasks, designed to assess different hand-eye co-ordination skills, were examined in a progressive manner, with the most complex task last (Table 1). The task settings were set to the default agreed by the MIST VR Expert User Group.

Repeat sessions were arranged so that subjects were on the trainer at similar times of day when they returned for follow-up sessions, as far as other commitments would allow, and tested in exam mode straight away.

At the end of each session, a print out of the subject's performance was analysed and sub-totals for time, errors, and efficient use of the diathermy pedal were entered onto an Excel spreadsheet which generated an 'overall score' for any subject's performance on any given task. The score is a function of total time taken, contact errors made and excess burn time on the diathermy pedal, and varies according to task. The formula by which the score was calculated was designed to penalise instrument errors more heavily than time taken to complete the individual tasks. For example, the formula used to derive the overall score for task 6 is $\{[(2a) + (4b^2) + (9c^2) + (d/2)]/10\}$, where a = excess time used for diathermy, b = total time movement outside bounding box, c = total number of contact errors and d = total time taken. The larger the overall numerical score, the worse the overall performance. Time taken for task completion and errors were also analysed separately to determine which contributed to the familiarisation process. Economy of hand movement was also correlated with overall score.

Statistical methods

The scores from the six tasks were all found to have positively skewed distributions. This was rectified by analysing the inverse of the scores. The transformed (inverse) scores from each task were analysed separately using one-factor repeated measures analyses of variance to compare mean scores across the six studies included in the analysis. Significant differences between consecutive studies were then evaluated by examination of the standard errors of the study scores differences. A P value of < 0.05 was taken as significant. Comparisons of changes in time taken and errors made were calculated using a paired t-test if the results were normally distributed, or a Wilcoxon's test if the data were non-parametric. The relationship between overall scores achieved on the first and third sessions, as well as between overall scores and economy of hand movement were analysed using Pearson's correlation, with two tailed significance.
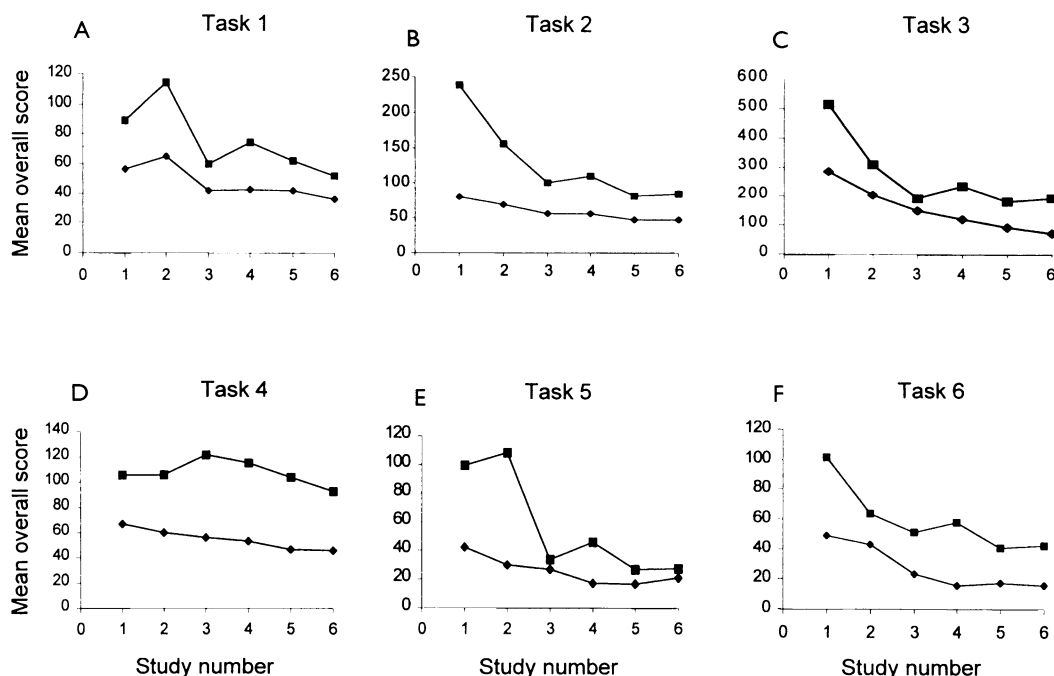
Figure 1 Mean overall score against study number for surgeons (filled diamonds) versus non-surgeons (filled squares) for tasks 1–6.

Table 2 Differences in mean scores and familiarisation curves between surgeons and non-surgeons

Task no	Difference in mean scores			Difference in familiarity curve		
.	f ratio	df	P value	f ratio	df	P value
1	4.8	1	0.044	0.192	5	0.965
2	10.7	1	0.005	0.032	5	0.999
3	3.8	1	0.069	1.786	5	0.126
4	11.53	1	0.004	1.119	5	0.322
5	5.96	1	0.027	1.97	5	0.092
6	9.23	1	0.008	2.97	5	0.017

Results

Data were available from all 18 subjects for five studies and for 16 of the subjects for a further sixth study. Thereafter, the number of subjects for whom analysis was available for 7–10 studies was too small, hence the statistical evaluation reported in this paper has been restricted to the first six studies.

Human-computer interaction and familiarisation

There was a clear and statistically significant reduction in mean scores between study number 1 and 3 within all 6 tasks (Fig. 1A–F). Changes in score after study 3

tended to be smaller and were non-significant, although a continued slight downward trend could be seen for most tasks.

There was no significant difference in the familiarity curves between surgeons and non-surgeons for all six tasks (Table 2). A visible difference in learning trend was seen in task 5 (Fig. 1E), although it was not statistically significant. The surgeons appeared to learn this task considerably quicker than the non-surgical group with no significant change in scores after the second session. The non-surgeons were slower to master this task, and their scores settled only after 4 sessions.

In terms of overall score, the surgical group scored significantly and consistently higher than the non-surgeons for all the tasks and for all the study sessions.

For task 5 (Fig. 1E) although non-surgeons lagged behind surgeons in terms of task/instrument familiarisation, by the sixth session overall scores achieved by the two groups were comparable. Our results suggest that this is the only MIST VR task where non-surgeon's performance improves to an equivalent standard as surgeons.

For task 6, arguably the most difficult combination of skills, the surgeons executed this task considerably better than the non-surgeons. In proportional terms, the surgeons improved their scores for this task by the

Table 3 Changes in errors made and time taken between study 1 and study 3 for the total population

Task no	Reduction in errors (P value)	Reduction in time taken (P value)
1	0.156	< 0.001
2	0.002	< 0.001
3	0.002	< 0.001
4	0.113	< 0.001
5	0.001	< 0.001
6	0.026	0.007

sixth session, by up to 70%, compared with the non-surgeons who managed to improve their scores by only 55–60%. This difference is statistically significant.

The improvement in overall score for all the tasks, was broken down into separate analyses of time taken and errors made, in order to examine how these two factors both improved. The results are summarised in (Table 3). There is a fall in errors made over all six tasks. For tasks 2, 3, 5 and 6, there is a statistically significant fall in errors made between session number one and three. For tasks 1–6 inclusive, there is a significant fall in time taken to complete the tasks between the first and third attempt. MIST VR measures economy of right and left hand movement as a ratio of the excess distance travelled by the instruments between object and target compared with the straight line trajectory. There was a significant correlation between overall score and economy of left-hand instrument movements for tasks 4 and 5, and economy with the right-hand for tasks 1, 5 and 6 (task 3 not analysed as not subject to handedness).

Predicting performance

Having clearly identified the presence of an early learning curve on the simulator for both surgeons and non-surgeons, the results were examined to see if it was possible to predict a subject's laparoscopic skills performance level on the simulator, from the results of their initial session. The results show that for every individual in the surgical and non-surgical groups, tasks 1–6 inclusive (Table 4), there was a statistically significant relationship between the overall score for the first and third studies for four out of six tasks, for the population of non-surgeons. As the number of non-surgeons was small, a further study was carried out using 17 final year medical students (aged 22–25 years). None of the students had any prior 'hands-on' experience of laparoscopic surgery although some had observed such procedures. They were each tested on the simulator according to the explicit protocol described previously. The same correlations between

Table 4 Spearman's correlation coefficients between overall score for study 1 and score for study 3

Population	Task no	Correlation coefficient	P value
Non-surgeons (n = 7)			
	1	0.0083	0.031
	2	0.2348	0.612
	3	0.517	0.235
	4	0.9978	< 0.0001
	5	0.8949	0.006
	6	0.9728	< 0.0001
Surgeons (n = 11)			
	1	0.0453	0.895
	2	0.455	0.894
	3	0.6125	0.045
	4	0.8892	< 0.0001
	5	0.5602	0.073
	6	0.4752	0.14
Medical students (n = 17)			
	1	0.7693	< 0.0001
	2	0.9772	< 0.0001
	3	0.5566	0.02
	4	0.7337	0.001
	5	0.5753	0.016
	6	0.5913	0.043

their overall score for the first session their scores for the third session were calculated (Table 4). There was a highly significant correlation between initial performance on MIST VR and the medical students' subsequent performance on session number three for all six tasks, once the familiarisation curve had been overcome.

Discussion

There is a clear need for a standard method of formally assessing a subject's psychomotor skills and laparoscopic surgical dexterity required for minimally invasive surgery, in an objective manner. MIST VR is a virtual reality simulator that 'abstracts' these skills and provides detailed analyses regarding an individual's performance, hence provides the basis for constructive feedback and rationalised training.

It has already been shown that MIST VR can objectively distinguish between experienced surgeons and non-surgeons.[10] In the same study, the author also reported the results of a small randomised controlled. In this trial using an alternative method of analysis to this present study, in order to assess performance level on MIST VR, trainee surgeons who had completed the basic surgical skills training course made fewer errors

and completed the tasks more efficiently than their untrained counterparts.

The present study has shown that the method of generating an overall score for performance (based on the computed formula for errors made and time taken) is a valid method of quantifying laparoscopic surgical experience. As expected, a population of surgeons scored consistently and significantly better than a surgically naive group.

This suggests that the scoring system and MIST VR are both sensitive and specific to measuring surgically relevant skills.

Importantly, this study identifies a rapid early 'learning curve' to the simulator. In effect, it takes any individual, surgically experienced or otherwise, up to three sessions to overcome the process of task learning and instrument/computer familiarisation. There was also a continued downward trend in score seen in most of the tasks after the third session on the simulator. These findings support a two-component hypothesis of initial learning followed by a period of skill refinement.[9] The most appropriate time to assess the subjects' underlying psychomotor skills will be when the 'familiarity curve' has plateaued.

This study has demonstrated the unexpected correlation between performance of surgically naïve individuals on their first session on MIST VR and their subsequent performance after overcoming the interface familiarisation curve. Thus, it is possible to predict how well a non-surgeon will perform on the simulator after overcoming the learning curve, without the requirement for repeated test sessions. This raises the possibility of using MIST VR as a 'one-off' aptitude test for individuals who are contemplating careers in minimally invasive surgery. Objective information regarding potential surgical ability could also be useful for juniors who are already on basic surgical training programs. However, before MIST VR could be used for this purpose, a large validated database of scores would be required to determine the 'entry' or 'pass' score. In the data already collected, it seems there are external factors that influence overall performance on the simulator. Work in progress examining both fixed and variable mental and physical parameters should help elucidate some of the additional factors affecting surgical performance as measured by MIST VR.

In conclusion, the issue of objective, comparative assessment of laparoscopic surgical competence needs to be addressed. Virtual reality simulators which mimic the movements required for laparoscopic surgery lead the way in providing non-subjective scores based upon efficiency and errors made. As in real surgical practise, a learning curve exists for the simulator, during which, the process of task, instrument and computer interface familiarisation is overcome and is important when evaluating initial performance on the simulator.

Acknowledgements

References

1. Shaper NJ, Harrison M, Bates T. Impact of laparoscopic cholecystectomy on surgical training. *Ann R Coll Surg Engl* 1996; **78**: 39–42

2. Wickham JE. Minimally invasive surgery. Future developments. *BMJ* 1994; **308**: 193–6

3. Department of Health. *Hospital Doctors: Training for the Future.* London: DoH, 1993

4. EAES guidelines. *Surg Endosc* 1994; **8**: 721–2

5. Sachdeva AK, Loiacono LA, Amiel GE, Blair PG, Friedman M, Roslyn JJ. Variability in the clinical skills of residents entering surgery. *Ann Surg* 1995; **118**: 300–8; discussion 308–9

6. Stone RJ, McCloy RF. Virtual environment training systems for laparoscopic surgery; activities at the UK's Wolfson Centre for Minimally Invasive Therapy. *J Med Virtual Reality* 1996; **1**: 42–51

7. Wilson MS, Middlebrook A, Sutton C, Stone R, McCloy RF. MIST VR: a virtual reality trainer for laparoscopic surgery assesses performance. *Ann R Coll Surg Engl* 1997; **79**: 403–4

8. Coleman J, Nduka CC, Darzi A. Virtual reality and laparoscopic surgery. *Br J Surg* 1994; **81**: 709–11

9. Mitta DA, Packebush SJ. Improving interface quality: an investigation of human-computer interaction and task learning. *Ergonomics* 1995; **38**: 1307–25

10. Taffinder N, Sutton C, Fishwick RJ, McManus IC, Darzi A. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery. In: Westwood RJ, Hoffman H, Stredney D, Weghorst S, eds. *Technology and informatics 50: proceedings of medicine meets virtual reality 6 – 1998 Jan 28–31,* San Diego, USA. Amsterdam: IOS Press, 1998: 124–30