The Royal College of Surgeons of England

REVIEW

Ann R Coll Surg Engl 2008; **90**: 282–285
doi 10.1308/003588408X286017

Assessment of surgical skills of trainees in the UK

JONATHAN D BEARD

Sheffield Vascular Institute, Northern General Hospital, Sheffield, UK

ABSTRACT

Surgical training and assessment in the UK has been criticised in the past for lacking transparency, reliability and validity. The new Intercollegiate Surgical Curriculum Programme (ISCP) has a well-defined, competence-based syllabus and a system of workplace-based assessments and examinations that map to the syllabus. The main aims of workplace-based assessment are to aid learning through objective feedback and to provide evidence that the competencies required to progress to the next level of training have been achieved. Reduction in surgical experience means that more training will need to be undertaken on simulations, although experience and assessment in the operating room must remains the 'gold-standard'. Simulation training will require the provision of properly resourced surgical skills facilities in every hospital. The key to reliable assessment and constructive feedback is well-trained trainers. Training is a skill that must be learned, and assessment and feedback techniques form part of this. In surgery, it has been assumed that all consultants are trainers but this is clearly not the case. Surgeons will need to follow the example of primary care, where trainers are selected from experienced general practitioners who demonstrate enthusiasm and ability. The reward for the trainer should be protected time for training. The reward for the National Health Service will be better trained surgeons.

KEYWORDS

Surgical skills – Trainees – Assessment – Intercollegiate Surgical Curriculum Programme

CORRESPONDENCE TO

Jonathan D Beard, Programme Director, Sheffield Vascular Institute, Northern General Hospital, Sheffield S5 7AU, UK
T: +44 (0)114 271 5534; F: +44 (0)114 271 4747; E: jonathan.d.beard@sth.nhs.uk

Surgical training in the UK has traditionally been based upon an apprenticeship and examination model. Trainees had to complete a set number of years of training and pass the Intercollegiate Fellowship of the Royal Colleges of Surgeons (FRCS), in order to achieve their Certificate of Completion of Specialist Training (CCST). Although consultant trainers completed a Trainee Assessment Form at the end of each placement, most received no training in assessment and were asked to make summary judgements based on little objective evidence. Unsurprisingly, trainers were reluctant to make 'negative' comments about trainees and the annual Record of In-Training Assessment (RITA) has fallen into disrepute because of the lack of evidence upon which to make a summary judgement. The examination, which consisted of a written paper, a clinical examination and vivas, is heavily knowledge-based and does not test technical skills. The reduction in the operative experience of trainees caused by shifts and limitations on hours of work also meant that operative competence could no longer be assured on the basis of experience alone.[1] This article explains some of the new assessments that are being introduced as part of the Intercollegiate Surgical Curriculum Programme (ISCP). This article does not address the issue of selection into surgery. Like examinations, selection is a 'high-stakes' assessment which is beyond the scope of this article.

Principles of assessment

Like all professionals, educationalists have their own jargon. This can be a cause of misunderstanding; therefore, some key terms are explained here. Assessment can be defined as making a judgement against a defined reference. This may be norm-referenced (*i.e.* judged against what is 'average', for example the old Trainee Assessment Form) or criterion-referenced against a defined statement, for example, 'can perform a particular operation without supervision'. A norm-referenced system implies that the 'average' level will increase with the seniority of the trainee. This level is rarely defined, as it depends upon the performance of the whole group, as well as the 'intuitive' judgement of the trainer. This partly explains the poor reliability of the Trainee Assessment Form, although reliability can be improved by describing positive and negative attributes.

Assessment has two main purposes, which ideally should not interfere with each other. The first is to provide feedback to aid learning (*i.e.* formative or low-stakes assessment) and the second is for examination/certification (*i.e.* summative or high-stakes assessment). The level of assessment can also vary: performance-based assessment measures what surgeons do in actual professional practice (*e.g.* measurement

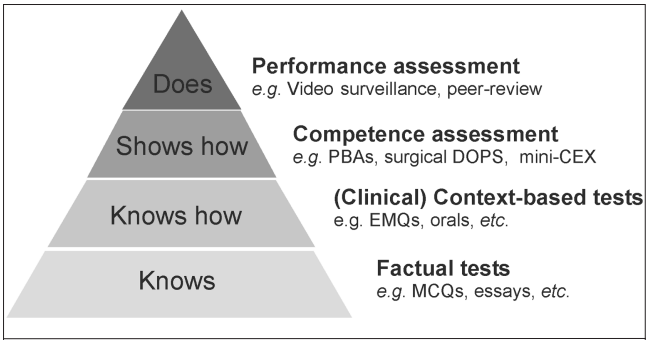

Figure 1 Miller's triangle defines a simple hierarchy for the development and assessment of clinical skills.

of outcomes). Performance can be affected by illness, attitude or environment which means that a competent surgeon may perform poorly. Competence-based assessment measures what a surgeon can do in a controlled representation of professional practice (*e.g.* when observed in the operating room or in a clinical skills laboratory).[2] Miller's triangle[3] defines a simple hierarchy for the development and assessment of clinical skills and the ISCP is based upon this model (Fig. 1). Some have suggested that there should be a higher level of 'does well' corresponding to an expert and that this, not mere competence, should be the goal of training.[4] 'Does well' probably requires more time for practice (experience) than can be achieved during training and indicates the need for continued learning after CCT.

A good assessment must possess reliability, validity, educational impact, acceptability and feasibility.[5] Reliability is a measure of reproducibility and discriminating ability. A reliable assessment should give the same result if repeated (test–retest) or if a different assessor is used (inter-rater). Reliability is an indication of these correlations and ranges from 0 (no reliability) to 1 (perfect reliability). An arbitrary, yet generally accepted, cut-off for a high-stakes assessment is 0.8.

Validity is an indication of how well an assessment measures what it claims to test. A valid assessment should contain the essential elements of the behaviour or skill being tested (content validity), experts should achieve higher scores (construct validity) and there should be agreement with other methods intended to measure the same thing (criterion validity). Ideally, the assessment should also predict the outcome (outcome validity), but this can be hard to measure. The study strategy of a learner is heavily influenced by assessment.[6] We tend to criticise trainees for this but it seems more sensible to capitalise on it by ensuring that the assessments map to the syllabus. This is called blueprinting, which ensures good educational impact. Finally, an assessment must be acceptable (covert surveillance may be reliable but it is probably unacceptable in most cases) and feasible (there is no point in developing a 'perfect' assessment that is too expensive to run).

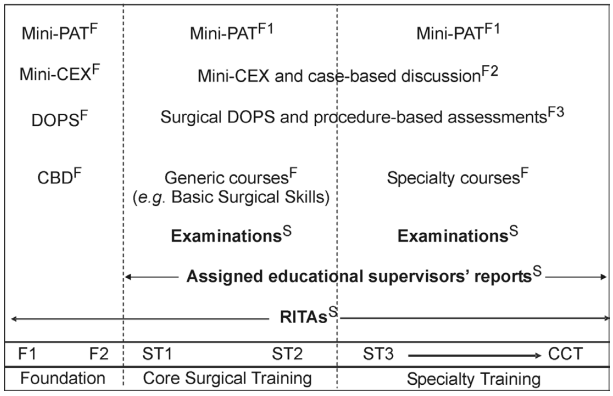

Figure 2 An outline of the ISCP assessment system. Workplace assessments (indicated F) are formative tools, designed to facilitate learning; they should be undertaken as often as possible. Summative assessments are indicated S. The content of assessments is generic in early training, becoming increasingly specialised as training progresses. (1) Mini-PAT is undertaken at entry to specialty training (for most in ST1) and every 3 years thereafter (for most in ST4 and ST7). It can be repeated if there is any concern about a trainee's professional behaviour. (2) Mini-CEX and CBD should each be undertaken at least 6 times per year in ST1/ST2. Their use in specialty training will depend on the specialty and level of training. (3) Surgical DOPS and PBAs facilitate the acquisition of surgical skills. Surgical DOPS should be undertaken at least 6 times per year in ST1/ST2. PBAs focus on index procedures in each specialty and should be used every time an index procedure is performed.

The other terms requiring definition are syllabus and curriculum. A syllabus consists of the core competencies (*i.e.* knowledge, skills, judgement and behaviours) required of a trainee. The curriculum defines when, where and how these will be learned, and also details the assessment system. This system has to conform to the principles that have been laid down by the Postgraduate Medical Education and Training Board (PMETB).[7] An outline of the ISCP assessment system is shown in Figure 2.

The future assessment of surgical skills

Many of the skills and professional behaviours required of a surgeon are the same as those required of all doctors, as laid out in the General Medical Council's *Good Medical Practice*.[8] These include knowledge, clinical skills, communication, team-working, judgement, compassion and probity. Assessment of these generic competencies by multiple raters (peer review) is reliable and well validated.[9] The Mini Peer Assessment Tool (Mini PAT) has been adopted by the Foundation Programme and the ISCP. The additional assessments required of surgeons are those dealing with surgical knowledge and skills. Surgical knowledge is probably best assessed by examinations but surgical skills are best assessed in the workplace.

Log-books form a useful record of experience gained,[10] but experience does not necessarily equate with competence. The OpComp form, developed and validated by the North Trent and South West Training Schemes, asked consultant trainers to assess the ability of a trainee to perform specific index operations against defined criteria at the end of a training placement.[11] This information complements the log-book and helps to identify deficiencies which may require targeted training during the next placement. Although an advance, the OpComp form suffers from poor reliability because the assessment is retrospective. Nor does it help trainees to learn 'on-the-job'.

Assessment in the operating room

It seems obvious that direct observation of surgical skills in the operating room represents the 'gold standard' in terms of content and construct validity. The technical competence of trainees in the operating room has been assessed using a two-part Structured Technical Skills Assessment Form (STSF) by Winckel *et al.*[12] Part 1 consists of the essential components of the procedure (Task Specific Checklist) and part 2 consists of more non-specific items, for example handling of instruments (Global Ratings). Such assessments possess good inter-rater reliability and have the added advantage of providing immediate feedback, which aids learning. Procedure-Based Assessments (PBAs) combine these two methods into a single form, as well as rating the ability of the trainee to perform the whole procedure. The form has been designed so that it can be quickly completed by a trainer and fed-back to the trainee between cases. PBAs have been adopted as the principal workplace-based assessment method for technical skills training by the ISCP. PBAs have been written for all index procedures in each surgical specialty and they can be downloaded from the ISCP website (<www.iscp.ac.uk>). It is envisaged that they will be used every time that a trainee performs an index procedure. At the end a training placement, the collection of PBAs, together with the log-book, will enable the Assigned Educational Supervisor to make a summary judgement about the competence of a trainee to perform a specific index procedure.

Whilst technical skills are vital, many adverse events are caused by failings in non-technical aspects; a recent study found that poor communication was a causal factor in 43% of errors in surgery.[13] Non-Technical Skills for Surgeons (NOTSS) assesses four categories of these high level skills – situation awareness, decision making, communication and teamwork, and leadership.[14] It provides a framework and common terminology that allows surgeons to communicate effectively with each other in this area of practice and will help trainees to develop these abilities in the workplace. A handbook can be downloaded from <www.abdn.ac.uk/iprcnotss>.

Video-recording of operations for subsequent analysis may prove useful when external assessment is required. A portfolio of recorded consultations forms part of the requirement for the Membership of the Royal College of General Practitioners.[15] Many operating rooms are now equipped with camera lights and video monitors. Beard *et al.*[16] has shown good inter-rater reliability between direct and video assessment of saphenofemoral disconnection. However, Scott *et al.*[17] found that global assessment of edited videotapes of laparoscopic cholecystectomies did not correlate well with direct observation. Reliability for more complex operations can be improved by voice recordings or other information, such as event-marking when a trainee requires instruction or help. Video-recordings complement PBAs and a combination of the two provides powerful feedback.

Assessment on simulations

Assessment may be easier to undertake on simulations in a skills centre than in the operating room, due to the non-threatening and controlled environment. Simple bench-top simulations and live animal simulations have been compared by Martin *et al.*[18] using Objective Assessment of Technical Skills (OSATS). Bench-top simulations gave equivalent results to the use of live animals. More recently, the same group have shown that the technical skills learned on simple bench-top simulations are transferable to a human cadaver model.[19] Beard *et al.*[16] and Datt *et al.*[20] have also shown that assessment on simple bench-top simulations, including knotting, suturing, vessel ligation and tissue dissection, predicts performance in the operating room. It seems possible, therefore, to deconstruct operations into their component parts. Trainees can practice on simulations representing each component, and be assessed as competent, before undertaking the complete operation.

The virtual reality laparoscopic simulator, developed by Taffinder *et al.*,[21] generates objective measurements of economy of motion, the number of movements made and the number of errors. These criteria have been validated for tasks in both reality and virtual reality. Scott *et al.*[22] found a significant correlation between skill testing on a virtual reality simulation and intra-operative assessment during laparoscopic cholecystectomy. One main limitation of virtual reality simulations has been the lack of a force-feedback system (haptics), *i.e.* no resistance is felt when an object is touched. Virtual reality simulators incorporating realistic 3-D graphics and haptic feedback are being developed, thanks to the rapid increases in the speed and capacity of personal computers, although they remain expensive. Automated assessment and feedback may reduce the need for direct observation by a trainer, although it is vital that these mirror reality: the assessment for an index procedure should also be applicable to a simulation.[23] It is unclear

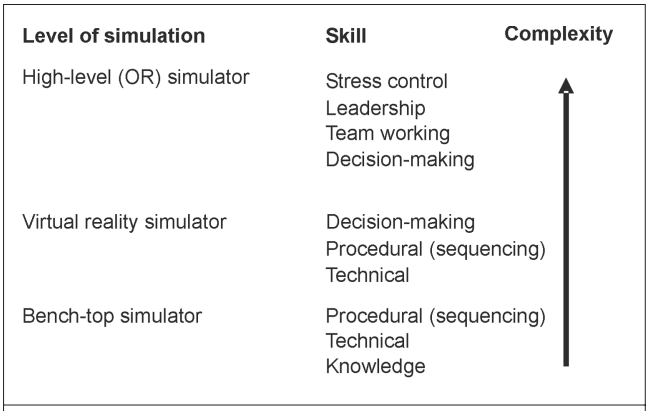

Figure 3 A hierarchy for simulation training.

whether virtual reality is better than simple bench-top simulation, although virtual reality does allow practice of the procedural steps and high-level virtual operating rooms can be used for untoward incident training of operating teams.[24] It seems logical that a trainee could progress up a hierarchy of simulators, as shown in Figure 3.

Issenberg *et al.*[25] suggested that simulations will form the basis for technical skills training and assessment in the future. This is because of the decreasing opportunity to practice on real patients and the need for 'deliberate practice' in a non-threatening environment.[26] Practice and assessment on simulations are no substitute for operative experience but they enable surgeons to become competent (and, therefore, confident) in key surgical skills before entering the complex operating room environment.[27] A randomised trial has shown that trainees who receive simulation training perform significantly better in the operating room.[28]

References

1. Lowry J, Cripps J. Results of the online EWTD trainee survey. *Ann R Coll Surg Engl Suppl* 2005; **87**: 86–7.
2. Rethans J-J, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T *et al*. The relationship between competence and performance: implications for assessing practice performance. *Med Educ* 2002; **36**: 901–9.
3. Miller GE. The assessment of clinical skills, competence and performance. *Acad Med Suppl* 1990; **65**: S63–7.
4. Talbot M. Monkey see, monkey do: a critique of the competency model in graduate medical education. *Med Educ* 2004; **38**: 587–92.
5. Schuwirth LWT. Assessing medical competence: finding the right answers. *Clin Teacher* 2004; **1**: 14–8.
6. Newble D, Jolly B, Wakeford R. *The certification and recertification of doctors: issues in the assessment of clinical competence*. Cambridge: Cambridge University Press, 1994.
7. Southgate L, Grant J. *Principles for an assessment system for postgraduate medical training*. London: Postgraduate Medical Education and Training Board, 2006 <www.pmetb.org.uk>.
8. General Medical Council. *Good Medical Practice*. London: GMC, 2006 <www.gmc-uk.org>.
9. Archer J, Norcini J, Southgate L, Heard S, Davies H. Mini-PAT (Peer Assessment Tool): a valid component of a National Assessment Programme in the UK? *Adv Health Sci Educ Theory Pract* 2008; In press.
10. Galasko C, Mackay C. Unsupervised surgical training: logbooks are essential for assessing progress. *BMJ* 1997; **315**: 1306–7.
11. Thornton M, Donlon M, Beard JD. The operative skills of higher surgical trainees: measuring competence rather than experience undertaken. *Ann R Coll Surg Engl* 2003; **85**: 190–3.
12. Winckel CP, Reznick RK, Cohen R, Taylor B. Reliability and constant validity of a structured technical skills assessment form. *Am J Surg* 1994; **167**: 423–7.
13. Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of error reported by surgeons at three teaching hospitals. *Surgery* 2003; **133**: 614–21.
14. Yule S, Flin R, Patterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. *Med Educ* 2006; **40**: 1098–104.
15. Joint Committee on Postgraduate Training for General Practice. Assessment. 2004 <www.jcptgp.org.uk/certification/assessment.asp>.
16. Beard JD, Jolly BC, Newble DI, Thomas WEG, Donnelly J, Southgate LJ. Assessing the technical skills of surgical trainees *Br J Surg* 2005; **92**: 778–82.
17. Scott DJ, Rege RV, Bergen PC, Guo WA, Laycock R, Tesfay ST *et al*. Measuring operative performance after laparoscopic skills training: edited videotape versus direct observation. *J Laparoendosc Adv Surg Tech* 2000; **10**: 183–90.
18. Martin JA, Regehr G, Reznick R, MacRoe H, Murnaghan J, Hutchinson C. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg* 1997; **84**: 273–8.
19. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M *et al*. Assessment of technical skills transfer from the bench training model to the human model. *Am J Surg* 1999; **76**: 167–70.
20. Datta V, Bann S, Beard JD, Mandalia M, Darzi A. Comparison of bench test evaluations of surgical skill with live operating performance assessments. *J Am Coll Surg* 2004, **199**: 603–6.
21. Taffinder N, McManus I, Jansen J, Russell R, Darzi A. An objective assessment of surgeon's psychomotor skills: validation of the MIST-VR laparoscopic simulator. *Br J Surg Suppl* 1998; **75**: 281.
22. Scott DJ, Bergan PC, Rege RV, Laycock R, Tesfay ST, Valentine RJ *et al*. Laparoscopic training on bench models: better and more cost-effective than operating room experience? *J Am Coll Surg* 2000; **191**: 272–83.
23. Gould DA, Kessel DO, Healey AE, Johnson SJ, Lewandowski WE. Simulators in catheter-based interventional radiology: training or computer games? *Clin Radiol* 2006; **61**: 556–61.
24. Kneebone R. Simulation in surgical training: educational issues and practical implications. *Med Educ* 2003; **37**: 267–77.
25. Issenberg SB, Gordon MS, Gordon DL, Safford RE, Hart I. Simulation and new learning technologies. *Med Teacher* 2001; **23**: 16–23.
26. Ericsson KA, Krampe RT, Tesch-Romer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993; **100**: 363–406.
27. Aggarwal R, Black SA, Hance JR, Darzi A, Cheshire NJW. Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *Eur J Vasc Endovasc Surg* 2006; **31**: 588–93.
28. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomised clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004; **91**: 146–50.