

ORIGINAL ARTICLE

## Evaluating trainee experience of surgical skills teaching

Simon Fleming,<sup>1</sup> Elizabeth Winson,<sup>2,\*</sup> Kurdo Amin<sup>3</sup> and Bijendra P. Patel<sup>4</sup>

<sup>1</sup>Whittington Hospital, Magdala Avenue, London N19 5NF, UK; <sup>2</sup>Launceston General Hospital, Launceston, Tasmania, Australia;

<sup>3</sup>Centre for Surgical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK; <sup>4</sup>Centre for Surgical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK

\*Corresponding author: Dr Elizabeth Winson, Launceston General Hospital, Launceston, TAS 7250, Australia. Email: emwinson@gmail.com

Date accepted for publication: 6 November 2014

### Abstract

**Introduction:** Operating theatres are a crucial learning environment for trainee surgeons developing surgical skills. There is no structured framework for teaching during surgery or its evaluation. Objective assessment of the learning experience can assess quality and highlight areas for improvement, maximizing benefit. **Methods:** A 5-point Likert rating scale was devised to assess surgical teaching experience. Positive and negative teaching attributes were established with a literature review and interviews of trainees and teachers. Sixty surgical trainees at a major London teaching hospital, operating under consultant supervision, evaluated the supervisor teaching using the tool. Significance of test results from the assessment tool was analysed using repeated measures analysis of variance for Likert scoring. A  $P$  value  $< 0.05$  was considered statistically significant. **Results:** There was consensus between trainees on the relevance of tool themes, with most ranked important or absolutely essential and no difference across grades ( $P \geq 0.05$ ). There was no difference in the overall supervisor scores between trainees with different levels of experience ( $P \geq 0.05$ ) or between procedures with different levels of complexity ( $P \geq 0.05$ ). Junior supervisors scored more highly overall than senior supervisors ( $P = 0.024$ ).

**Discussion:** The study demonstrated that the assessment tool is feasible, practical and applicable, with face and content validity. Discrimination between supervisors with different levels of experience shows discriminative validity. The tool facilitates objective assessment of the teaching experience in surgery. We hope this will aid improvements in teaching quality, an area for further study.

**Keywords:** simulation; surgical training; evaluation

### Introduction

The operating theatre is the most crucial learning environment for trainee surgeons as they develop their surgical skills. Learning in theatre is opportunistic and relies on exposure to cases. It is limited by the complexity of the cases available and training time. Historically, training in this arena has followed an apprenticeship format, with little formal training or regulation of the trainers.<sup>1</sup>

In recent years, approaches to postgraduate medical education have changed, as illustrated by the Modernising Medical Careers (MMC) national reform in the United Kingdom. There are increasing efforts to make the careers of junior doctors more efficient, with the acquisition of key competencies and early identification of weaknesses.<sup>2</sup>

However, obstacles remain to the adequate education of surgical trainees. Despite recognition of the need to teach the teachers, there are many supervisors whose teaching skills are based solely on their own experience as surgical juniors.<sup>3</sup> In addition, the time available for training has been markedly compromised, for example with the introduction of the Working Time Directive in the United Kingdom.<sup>4</sup>

There is dissatisfaction apparent in both trainees and trainers with the current limitations on surgical training.<sup>5</sup> It seems clear that, under such constraints, there is a pressing need to ensure that all training delivered is of the highest standard. Despite this, there is no structured framework for teaching during surgery or direct assessment of teaching quality as measured by skills acquisition or trainee satisfaction. Objective assessment of trainees' teaching experience

can evaluate trainer performance and highlight areas for improvement, maximizing benefit to trainee and teacher from this limited resource.

Evaluation of teaching quality should therefore be a key responsibility of NHS trusts. Multiple sources of information can play a role in evaluation, but feedback from trainees is a useful resource to directly reflect the trainee experience. Trainee questionnaires provide an easily available evaluation tool that can provide quantifiable data. They can be tailored according to the specific attributes being assessed.<sup>6</sup>

Current tools in use to assess trainee acquisition of surgical skills, such as the Direct Observation of Procedural Skills (DOPS) method, focus directly on trainee performance of skills. Whilst repeated assessment of trainee performance over a period of teaching by supervisors may encourage inferences about the effectiveness of this teaching, the information gained will be indirect.

There is no tool in use at present designed to directly assess the ability of consultants as trainers in the operating theatre. A tool assessing trainee experience of consultant trainers would compliment existing information on skills acquisition from tools such as the DOPS, enabling broader assessment of teaching efficacy.

Validity requires a tool to be credible, appropriate and able to measure what it claims to measure. A range of different criteria exist to test the validity of an assessment tool. Construct validity requires that a test measures what it proposes to measure. Face validity estimates whether a test looks valid to its participants, whilst content validity estimates whether a measure represents each element of the relevant construct.

This study aimed to create an instrument with which to assess surgical supervisors' teaching of surgical skills, including in the peri-operative period, from the trainees' perspective. The tool was designed to focus on trainee satisfaction with important educator attributes of surgical supervisors. The study also demonstrates the validity of the tool in assessing the experience of surgical teaching, with a view to presenting it as a useful tool to evaluate the teaching of surgical technical skills.

## Method

A modified 5-point Likert rating scale was developed for trainees to assess supervisor teaching. The tool was based on a scale developed at St Mary's Hospital, which had already been shown to have face, content, concurrent, construct and predictive validity.<sup>7</sup> This was modified to include assessment of teaching behaviours. Positive and negative

attributes were ascertained based on interviews with trainees and trainers and a literature review.

The modified scale assesses supervisor performance in four domains: Briefing, Surgery, Debriefing and Behaviour, with grading from 1 (worst) to 5 (best) (Fig. 1). In addition, trainees were asked to rank the importance and relevance of the components assessed.

Trainees were subdivided according to experience as Senior House Officer (SHO) year 1–3, SHO year 4–6, Specialist Registrar (SpR) year 1–3 and SpR year 4–7. Trainees with 4 or less years of surgical experience were classified as junior.

Surgical procedures were defined by complexity and skill level as minor (e.g. excision of a lump), intermediate (e.g. inguinal hernia repair) and major (e.g. nephrectomy). In addition, supervisors were categorized into junior trainers if they had been consultants for less than 5 years and senior trainers if they had more than 5 years' such experience.

Sixty surgical trainees at St Bartholomew's and the Royal London who completed surgical procedures supervised by consultants were asked to evaluate the quality of supervisor teaching at the end of the day's theatre list. Half were junior trainees and half senior trainees. They were interviewed to ensure they were surgical trainees at ST1 level or above who had performed at least part of the procedure under consultant supervision. The aim of the study was explained to them and they were then asked to complete the questionnaire. The questionnaire was found simple to use by trainees and took an average of 15 minutes per trainee.

Statistical significance of results yielded from the assessment tool as used by trainees of different grades and for procedures of different complexities was analysed using repeated measures analysis of variance (ANOVA) for Likert scoring as global scores were parametric and more than two groups were compared. A *P* value <0.05 was considered to be statistically significant. Discriminatory validity of the tool for trainers was assessed by discriminating between junior and senior trainers and applying ANOVA as the global scores were parametric and two categories were being compared.

## Results

The results are divided into four categories: assessing the validity of the tool at different grades of trainee; assessing validity for different skill levels of surgical procedure; comparing junior and senior trainees; and assessing the relevance of themes in the tool. Sixty trainee surgeons completed the scoring questionnaire.

Theme	Skill	1	2	3	4	5
<b>Briefing</b>	Trainee awareness	Nil checks		Adequately explored		Well established
	Explanation	None		Adequate explanation relevant aspects of surgery		Clear explanation
	Preparation	Nil explanation / demonstration		Adequate		Clear and detailed explanation / demonstration
	Aims	No explanation		Adequately outlined		Clear and detailed
	Trainer	Nil guidance		Reasonable guidance when will take over		Precise guidance
<b>Surgery</b>	Assistance	Minimal assistance, poor concentration		Adequate assistance		Excellent and helpful assistance
	Time	Pressure on trainee to hurry / time constraints		Reasonable pressure to hurry / time constraints		No pressure to hurry / time constraints
	Errors	Does not allow trainee to continue after 1 minor technical error		Allows trainee to continue after a few minor technical errors		Allows trainee to continue after several minor technical errors or 1 major error which does not affect patient safety
	Difficult situations	Immediately takes over		Takes over after short trainee trial		Tries to take trainee through patiently before taking over
	Handing back	Does not		After few major steps		Immediately
	Explanation & demo post mistake	None		Adequate		Excellent, clear and detailed
	Demeaning trainee	Always		Occasionally		Never
	Surgical field	Trainer constantly interfering in surgical field		Occasionally interfered		Does not interfere in surgical field
	Autonomy	Dictatorial supervision, negative comments		Some autonomy		Complete autonomy, positive encouragement
	Explanation of technical tasks	Nil		Adequate		Clear
	Demo of technical tasks	Nil helpful		Adequate		Clear demo without interference
	Attention to trainee point of view	Never		Occasionally		Always
	Writing op note	By supervisor; no trainee teaching/input		Occasional review/input for trainee		Supervision/teaching whilst trainee writes
	<b>Debrief</b>	Feedback	Nil		Adequate feedback on surgical strengths/weaknesses	
Future		No future outline		Adequately outlines future surgical goals and objectives		Clearly outlines future goals and objectives
Confidence gained		Lose confidence		Some		Very high
Supervision beneficial		No		Good for routine procedures		Excellent for any procedure
<b>Behaviour</b>	Allows trainee communication and leadership	Never		Occasionally		Always
	Commanding or fierce	Always		Occasionally		Never
	Interrogating	Always		Occasionally		Never
	Hostility or frustration	Always		Occasionally		Never
	Encouraging	Never		Occasionally		Always

Figure 1 Modified 5-point Likert rating scale (abridged).

There was no significant difference in the overall score awarded to the supervisor by trainees with different levels of experience ( $P \geq 0.05$ ) or in any of the individual domains: briefing ( $P \geq 0.05$ ); surgery ( $P \geq 0.05$ ), debriefing ( $P \geq 0.05$ ) and behaviour ( $P \geq 0.05$ ) (Fig. 2).

No difference was found between the total scores awarded to supervisors for procedures of different complexities or in any of the individual domains ( $P \geq 0.05$  for all) (Fig. 3).

A clear difference was demonstrated between supervisors with different levels of experience, in keeping with previous reports.<sup>8</sup> Junior supervisors scored significantly more than senior supervisors for total scores ( $P = 0.024$ ) and in the surgery domain ( $P = 0.030$ ) (Fig. 4). There was a trend towards higher scores in junior trainers compared to senior trainers in the other domains, however this was not statistically significant.

Trainees demonstrated consensus on the relevance of themes in the scale with the majority ranked as important or absolutely essential and no difference across grades ( $P \geq 0.05$ ).

## Discussion

It seems clear that in health systems with significantly reduced training hours, surgical trainees must receive an ever higher standard of surgical training and supervision in order to ensure the development of competent surgeons. Despite this, formal assessment of supervision and training is rare in the most important learning environment of all, the operating theatre.

A simple, well validated tool to assess the teaching attributes of consultants supervising surgical procedures provides an opportunity to evaluate and ultimately improve these learning experiences.

The tool developed for this study aimed to assess the teaching attributes of consultant trainers, as experienced by their trainees. The tool was developed to include attributes deemed important by trainers and trainees alike, taking into account the available literature. Development of the tool through detailed interviews with relevant parties and a review of the literature gave it face and content validity. Trainee consensus was obtained for all themes in the tool for their relevance and importance, with the majority of the themes ranked as absolutely essential or important.

Some attributes assessed by the tool are undoubtedly subjective and may vary depending on trainee personality, such as whether a trainer was demeaning or hostile. Despite this

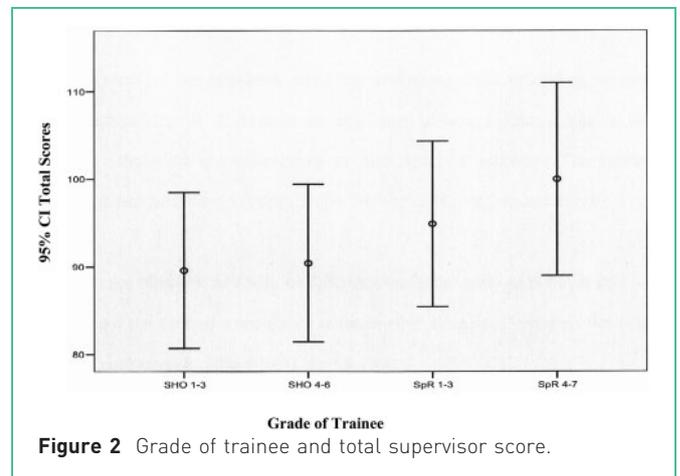


Figure 2 Grade of trainee and total supervisor score.

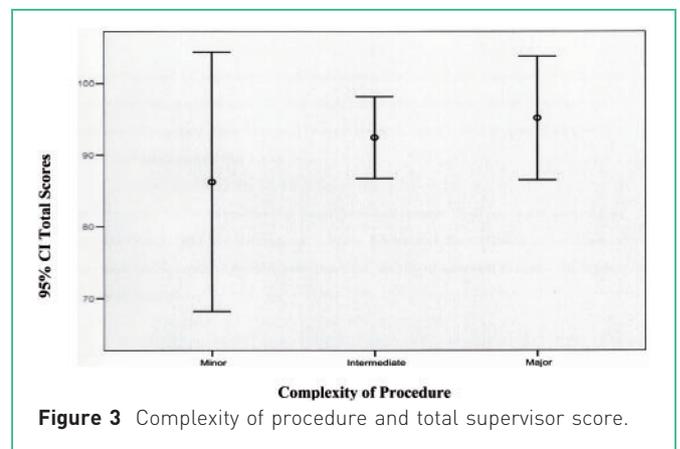


Figure 3 Complexity of procedure and total supervisor score.

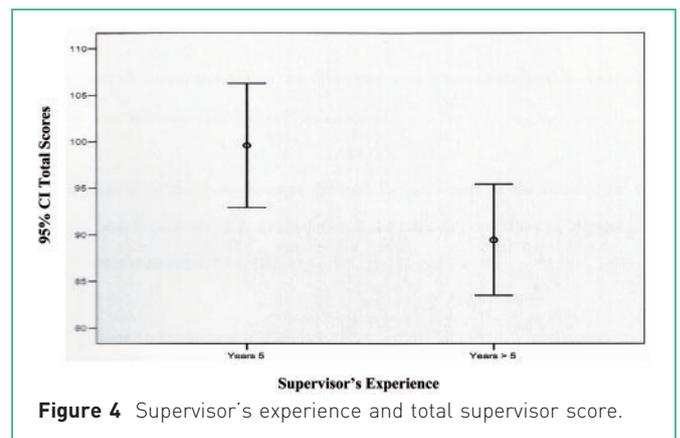


Figure 4 Supervisor's experience and total supervisor score.

we believe that assessment of these subjective human factors is of value, as repeated identification of attributes by different trainees can guide trainer reflection and improvement. If multiple trainees find a trainer attribute detrimental to their teaching, this warrants identification.

Our study demonstrated that the assessment tool is feasible, practical and easily applicable across a range of trainees performing various procedures.

In keeping with a previously published report,<sup>8</sup> using our tool to compare teaching performance between junior and senior supervisors showed clear discrimination. This demonstrates predictive and discriminative validity.

More work is needed to explain the favourable assessment of more junior supervisors. However we can infer that this may reflect the recent shift towards training the teachers in educational delivery and a tendency for younger consultants to be perceived as less intimidating or didactic.

The main limitation of the study is that we have not conducted a full assessment of the reliability of the tool. We have assumed reliability based on previous pilot studies.<sup>7,8</sup> Future work should explore the reliability of the assessment tool by comparing the trainee's feedback with observational data, for example, using blind video recording. In addition, further work should seek to establish a link between trainee satisfaction with the learning experience and actual, quantifiable skill acquisition.

Nonetheless, the assessment tool can play a role in the objective assessment of teaching and learning in surgery. This can provide feedback to supervisors, illustrating individual areas for performance enhancement. In addition, the tool can be used at a departmental level, allowing identification and dissemination of the qualities of good teachers. As an electronic version of the tool has now been devised, this can be extended to national level to assess consultant teaching across the country.

## Conflict of interest

No conflicts of interest have been declared.

## References

1. Macintyre IM. UK surgical training: current problems and possible solutions. *J R Coll Surg Edinb* 1996; 41: 209–12.
2. Audrey Ng, Gira P. MMC: must medicine change? *BMJ Career Focus* 2006; 333: 214–15.
3. Calman K. Hospital doctors: training for the future. *Br J Obstet Gynaecol* 1995; 102: 354–6. doi: 10.1111/j.1471-0528.1995.tb11283.x.
4. The surgical workforce in the new NHS. A report prepared by the Royal College of Surgeons of England. November 2001. Available at [http://www.rcseng.ac.uk/publications/docs/surgical\\_workforce.html](http://www.rcseng.ac.uk/publications/docs/surgical_workforce.html). Accessed October 2012.
5. Hindle KS, Carney L. Re: Registrar operating experience over a 15 year period: more, less or more or less the same? *Surg JR Coll Edinb, Irel* 2004; (2)3: 161–64. *Surgeon* 2005; 3: 56–7. doi: 10.1016/S1479-666X(05)80020-4.
6. Fry H, Kettering S, Marshall S, editors. A handbook for teaching and learning in higher education. 2nd edn. London: Routledge Falmer Publishing; 2003.
7. Sarker SK, Vincent C, Darzi AW. Assessing the teaching of technical skills. *Am J Surg* 2005; 189: 416–8. doi: 10.1016/j.amjsurg.2004.09.016.
8. Sarker SK, Amygdalos I, Darzi R, Kneebone R. Assessing the teaching of technical skills in operations and procedures. *Br J Surg* 2006; 93: s53–54. doi: 10.1002/bjs.5436.