Development of a rating system for surgeons' non-technical skills

S Yule,[1] R Flin,[1] S Paterson-Brown,[2] N Maran[3] & D Rowley[4]

BACKGROUND Analyses of adverse events in surgery reveal that many underlying causes are behavioural, such as communication failure, rather than technical. Non-technical (i.e. cognitive and interpersonal) skills are not addressed explicitly in surgical training. However, surgeons need to demonstrate these skills, which underpin their technical excellence, to maximise patient safety in the operating theatre. This paper describes the method used to identify surgeons' non-technical skills, and the development of a skills taxonomy and behavioural rating system to structure observation and feedback in surgical training.

METHODS Cognitive task analyses (critical incident interviews) were conducted with 27 consultant surgeons in general, cardiac and orthopaedic surgery. The interviews were coded and a multidisciplinary group of surgeons and psychologists used an iterative process to develop a skills taxonomy. This was supported by data gathered from an attitude survey, literature review, analysis of surgical mortality reports and observations in theatre.

RESULTS Five categories of non-technical skills were identified, including situation awareness, decision making, task management, leadership and communication and teamwork. This provided a structure for a prototype skill taxonomy (v1.1), which comprised 14 non-technical skill elements. Observable behaviours (markers) indicative of good and poor performance were developed for each element by 16 consultant surgeons to form a prototype behaviour rating system.

CONCLUSIONS The prototype skills taxonomy and behaviour rating system are grounded empirically in surgery. The reliability of the system is currently being tested using standardised scenarios. If this evaluation proves successful, the system could be used to structure feedback and guide non-technical skills training.

KEYWORDS surgery/*education; professional practice/*standards; task performance and analysis; interprofessional relations; consultants; Scotland; medical staff, hospital/*standards; multicenter study [publication type].

[1]School of Psychology, University of Aberdeen, UK
Departments of [2]Surgery and [3]Anaesthesia, Royal Infirmary, Edinburgh, UK
[4]Department of Orthopaedic and Trauma Surgery, University of Dundee, UK

Correspondence: Dr Steven Yule, School of Psychology, University of Aberdeen, Aberdeen AB24 2UB, Scotland, UK. Tel: +44 1224273214; Fax: +44 1224273211; E-mail: s.j.yule@abdn.ac.uk

INTRODUCTION

Analyses of adverse events in health care have revealed that many underlying causes originate from failings in non-technical aspects of performance rather than a lack of technical expertise.[1,2] In a recent study, communication was found to be a causal factor in 43% of errors made in surgery.[3] These findings support the argument that technical skills are necessary but not sufficient to maintain high levels of patient safety over time. In order to achieve this, attention needs to be paid to non-technical skills defined as behavioural aspects of performance in the operating theatre which are not related directly to medical expertise, use of equipment or drugs. These include the cognitive and social skills which underpin clinical and technical skills and have been identified as requirements for a competent surgeon.[4] Behavioural marker systems are already used to structure training and evaluation of these skills in anaesthesia[5]

Overview

What is already known on this subject

Failures in surgeons' non-technical skills (e.g. awareness, communication, teamwork) play a contributory role in surgical adverse events but surgical education focuses on training and assessing only technical skills.

What this study adds

A task analysis of surgeons' non-technical skills was conducted to determine which skills are important, before they can be trained and assessed.

A behaviour rating system was then developed to allow surgeons to structure observations, rate skills and provide feedback during a postoperative debrief.

Suggestions for further research

Future research will establish the criterion validity of non-technical skills ratings on performance in the operating theatre and on patient safety outcomes.

(as well as other 'high-risk' professions such as civil aviation and nuclear power). These marker systems are rating scales based on skills taxonomies and are used to evaluate observable behaviours that underlie performance. As with any assessment system, the behaviour rating tool needs to be explicit, transparent, reliable and valid for effective assessment of non-technical skills.[6]

To date, the formal training of surgeons has focused predominantly on developing knowledge, clinical expertise and technical skills. Aspects of performance such as decision making, leadership and team working have been developed in an informal and tacit manner rather than being addressed explicitly in training. Although there are existing checklists which cover some aspects of surgeons' non-technical ability[7–9] and a new focus on surgical competence,[10] there is currently no systematically developed system to rate behaviours during the intraoperative phase of surgery. Once

developed, such a tool could be used to structure observations, ratings and feedback in theatre as well as identify training needs and form the basis for non-technical skills training. We acknowledge that the intraoperative phase is only 1 component of the surgeon's task and that aspects of perioperative care have an impact on non-technical skills in the operating theatre. However, the scope of the present project was focused tightly on performance during the intraoperative phase of surgery and guided by the following research question: 'What are the main non-technical skills required by surgeons for safe surgical practice, and what observable behaviours are associated with those skills?'.

Behavioural marker systems

Behavioural marker systems are used to structure observation and evaluation of non-technical skills in high-demand professions[11] in order to improve safety and efficiency. They are context-specific and must be developed for the domain in which they are to be used. The systems developed for pilots[12] and anaesthetists[13,14] were developed systematically and subjected to experimental and practical evaluation. The aim of behavioural marker systems is to allow raters to identify 'observable, non-technical behaviours that contribute to superior or substandard performance'[15] (p. 10). Behavioural marker systems tend to comprise two parts: a skills taxonomy, with a set of behavioural markers allied to each skill, and a rating system. The skills taxonomy is usually developed on the basis of cognitive task analysis with domain experts in conjunction with analyses of adverse events, a literature search and traditional task analysis. These methods are used to identify the non-technical skills that are required for the profession in a given context (e.g. the operating theatre, flight deck). Once a skills taxonomy has been identified, behaviours indicative of good and poor practice are written. Domain experts normally provide these examples and a systematic method of writing and refining the indicative general behaviours is used.

This paper outlines the development of a skills taxonomy and behavioural rating system, specifically for individual surgeons' intraoperative non-technical skills. The taxonomy was derived from cognitive task analysis, supported by other methods, including a literature review on surgeons' non-technical skills[16] and a survey of theatre personnel attitudes to teamwork, error and safety.[17] Gordon's model of systems design[18] was used to guide the iterative

development of the system. This 3-phase model maps the process from task analysis through system design to evaluation. Only the interviews are discussed in detail in the present paper.

METHODS

Consultant surgeons ($n = 27$) from 11 hospitals in Scotland were recruited to participate in the study after being contacted by letter. They were interviewed at their hospital using the critical incident technique,[19] a type of cognitive interview.[20] The purpose of the interviews was to identify key non-technical skills from the discussion of critical incidents in the operating theatre. Cognitive interviews are particularly well suited for eliciting tacit knowledge in domain experts. Focusing on a specific memorable incident provides insight into the surgeon's use of information, strategies, meta-cognition and resources[21] during the intra-operative situation in an operating theatre. Tacit knowledge is difficult to gather using other methods[22] (for example, verbal protocol analysis could be used, but would probably interfere with the task and does not allow for reflection and inference to be made). In addition to insight in cognitive skills, the interview was also designed to uncover the interpersonal skills that were being used by the surgeon during the case.

Interview protocol

A multidisciplinary steering group comprising psychologists, surgeons and an anaesthetist developed the interview schedule (see Appendix 1). Surgeons were asked to recall events in theatre during a challenging, non-routine case and were probed about the course of events a further 2 times. After the surgeon described the case, the interviewer recounted the sequence of events back to the surgeon and asked for clarification and more explanation of the course of events. This second sweep of the case allowed for more details to be gleaned. The case was then discussed for a third time, with the addition of cognitive cues which recreate aspects of the case to elicit deeper-held tacit knowledge about the non-technical skills that were or were not being used. Examples of the cognitive cues used include: 'what cues were you using to help understand the situation' and 'how did you re-establish goals?'. A pilot study was conducted with 3 consultant surgeons to establish the face validity and utility of the technique. Minor changes to the method were made on the basis of results from the pilot study.

Sample

The specialist areas of the surgeons were general surgery ($n = 13$), orthopaedic surgery ($n = 10$) and cardiac surgery ($n = 4$). One of the participants was female. Ethical approval was granted by the Ethics Committee at the School of Psychology, University of Aberdeen.

RESULTS

The interviews lasted between 45 min and 1 hour. A variety of cases were discussed, including emergencies with duodenal ulcers, difficulties in hip and knee replacements, problems in transplant operations and difficulties with cardiac bypass.

Analysis of interviews

The interviews were recorded digitally, transcribed verbatim, and analysed using the line-by-line coding technique from grounded theory. This is an inductive technique that is suitable for exploring data and building theory[23] in order to aid system development. Coders were asked to identify when non-technical skills were discussed in the interview and to interpret those specific skills. They were provided with the definition of surgeons' non-technical skills (see Introduction) and an example-coded section of transcript. Three pairs of psychologists who were experienced at coding interview transcripts each coded 6 transcripts independently to an acceptable level of inter-rater reliability. The remaining transcripts were then coded. This process produced a list of 150 unsorted non-technical skills.

System development

The goal was to develop a system that could be used by surgeons in theatre rather than an instrument for research or an all-encompassing taxonomy. The prototype system was developed in 3 phases, and the tri-level hierarchical format used for behavioural marker systems in anaesthesia[24] and European civil aviation[25] was adopted. This format structures skills into category and element levels with observable behaviours (markers) indicative of good and poor performance for each element.

The aim of phase 1 was to develop a skills taxonomy, comprising category and element levels, which would form the basis of the system. To achieve this, the multidisciplinary research group reduced and refined the list of 150 skills extracted from the

transcripts. At this stage, the results of the literature review, survey and observations in theatre were also considered. The skills taxonomy was developed according to design criteria derived from the Joint Aviation Requirements: Translation and Elaboration of Legislation (JARTEL) project,[12] an expert panel on behavioural markers[15] and from cognitive task analysis[26] (see Table 1).

The reduced skills list was then organised thematically and broad categories emerged with component elements. In phase 2, an iterative process was used with 4 independent panels of consultant surgeons from 4 hospitals, who modified the structure into a skills taxonomy (v1.1) comprising 5 categories of behaviour and 14 skill elements (see Table 2). The panels checked the wording and labelling of elements, and ensured that the framework was relevant to the surgical domain. This formed the basis for the behavioural marker system.

In phase 3, observable behaviours (markers) indicative of good and poor performance were developed for each element by 16 consultant surgeons. The surgeons were asked to think of behaviours that could be either observed directly or inferred through communication. Two subsequent multidisciplinary review meetings refined this set of illustrative behaviours, all phrased as active verbs. This ensured that the system had cognitive and interpersonal functionality, was grounded in surgery and complied with the guidelines on system design[18] and behavioural markers design criteria[15,26] outlined above. Table 3 provides an example of the tri-level system for the category 'situation awareness'.

The NOTSS rating scale

A number of different scale formats and length can be used to rate observed behaviour.[27] A 4-point scale was chosen to allow the observable skills to be rated,

Table 2 NOTSS skills taxonomy v1.1

Category	Element
Situation awareness	Gathering information
	Understanding information
	Projecting and anticipating future state
Decision making	Considering options
	Selecting and communicating option
	Implementing and reviewing decisions
Task management	Planning and preparation
	Flexibility/responding to change
Leadership	Setting and maintaining standards
	Supporting others
	Coping with pressure
Communication and teamwork	Exchanging information
	Establishing a shared understanding
	Co-ordinating team activities

as follows: 4 good, 3 acceptable, 2 marginal, 1 poor, and NO not observed. The scale format allows for a more finely-grained analysis of performance than 2- or 3-point formats and is compatible with the standard method of surgeons' assessments in the United Kingdom. It is also compatible with other behaviour rating systems developed for use in acute medicine such as ANTS (anaesthetists non-technical skills).[24] The NO rating applies when the behaviour has not been demonstrated because there was no requirement for that skill to be used in the given case or scenario. If the skill should be observed but is not, then a rating of 1 (poor) should be given. Behaviours which potentially endanger patient safety should also be rated as 1 (poor). Following experience with the ANTS rating system, it is suggested to raters that they may find it more meaningful to rate the elements before rating the global categories.[28]

DISCUSSION

The prototype NOTSS behavioural marker system was designed to allow structured observation and

Table 1 Design criteria for behavioural markers, derived from previous work in the field[12,20,21]

1. The skills must be applicable to a surgeon's behaviour during the intraoperative phase of an operation
2. The system should comprise specific, observable behaviours that are well defined and contribute to superior or substandard performance
3. The skills and behavioural markers should either be directly observable in the case of social skills or inferred from observing communication or other behaviours, in the case of the cognitive skills
4. The system should be parsimonious and encompass the most important behaviours in the least number of categories and elements possible
5. The rating tool will need to fit onto 1 page, not larger than A4 paper, to be of practical use in the operating theatre or high-fidelity simulated environment. This will limit the number of categories and elements
6. The categories and elements should have the maximum mutual exclusivity possible. It is understood that this is achievable only to a certain degree, given the interdependence of the non-technical skills
7. The terminology used should reflect everyday, domain-specific language for surgeons' behaviour, rather than psychological jargon

Table 3 Prototype behavioural marker system for 'situation awareness' (note that only a sample of good and poor behaviours are provided for illustrative purposes. For the complete set of behaviours please visit http://www.abdn.ac.uk/iprc/notss)

Category	Element	Good behaviours	Poor behaviours
Situation awareness	Gathering information	Ensures that all relevant investigations (e.g. imaging) have been reviewed and are available	Arrives in theatre late or has to be repeatedly called
		Liaises with anaesthetist regarding anaesthetic plan for patient	Does not ask for results until the last minute or not at all
	Understanding information	Looks at CT scan and points out relevant area	Overlooks or ignores important results
		Reflects and discusses significance of information	Asks questions which demonstrate lack of understanding
	Projecting and anticipating future state	Plans operating list taking into account potential delays due to surgical or anaesthetic challenges	Gets into predictable blood loss, then tells anaesthetist
		Verbalises what may be required later in operation	Operates beyond level of experience

feedback to trainee and consultant surgeons on observable non-technical skills. A hierarchical structure of categories, elements and behaviours was adopted and 5 categories comprising 14 element skills emerged from the cognitive task analysis and other data gathered from surgeons. The identification of these non-technical skills will not come as a great surprise, as good surgeons have always displayed these behaviours, and the set of skills identified correlate well with previous studies looking at competence requirements for surgeons.[4] However, this study has taken the additional step of identifying explicitly the particular non-technical skills which might be both taught and observed – and therefore measured. This behavioural marker system now requires to be evaluated and, to this end, we are currently using standardised scenarios filmed in a simulated theatre environment, which are being analysed using the rating system by consultant surgeons. If the system can be shown to be both reliable and reproducible during this experimental evaluation, then the next step will be to trial it in an operating theatre environment.

Subject to these evaluations proving successful it is proposed that, with appropriate training in the use of the system, clinicians will be able to rate performance on these skills using the 4-point scale and provide feedback on behaviours in a systematic manner. By so doing, not only will it be possible to observe and assess non-technical skills in a formative manner, but also enhance training in this vital area of surgical performance. The Royal Colleges of Surgeons of England has recently adopted crew resource management (CRM) training and the Royal College of Surgeons of Edinburgh began running training

courses in non-technical skills for surgeons in early 2006. In addition to the practical uses of the system, the development of a standardised way to observe surgeons' intra-operative performance opens up a number of research applications aimed at improving levels of performance and safety in the operating theatre.[29]

CONCLUSIONS

The NOTSS skills taxonomy and marker system presented here has been grounded empirically in surgery, and was developed with domain experts (consultant surgeons) at every stage to ensure that the system is explicit, transparent and has an acceptable degree of construct validity. The next step is to test the reliability and usability of this system and its potential to guide the development of non-technical skills training for surgeons, much as is current practice in UK civil aviation where pilots' non-technical skills are assessed annually in licence revalidations.[30]

Ethical approval: ethical approval was granted by the School of Psychology, University of Aberdeen Ethics Committee.

REFERENCES

1 Bogner M, ed. *Human Error in Medicine.* Hillsdale, NJ: LEA 1994.

2 Bogner M, ed. *Misadventures in Health Care.* Mahwah, NJ: LEA 2004.

3 Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. *Surgery* 2003;**133**:614–21.

4 Baldwin PJ, Paisley AM, Brown SP. Consultant surgeons' opinion of the skills required of basic surgical trainees. *Br J Surg* 1999;**86**(8):1078–82.

5 Fletcher GCL, McGeorge P, Flin RH, Glavin RJ, Maran NJ. The role of non-technical skills in anaesthesia: a review of current literature. *Br J Anaesth* 2002;**88**(3):418–29.

6 Nunnally J, Bernstein I. *Psychometric Theory,* 3rd edn. New York: McGraw-Hill 1993.

7 Helmreich RL, Schaefer HG. Team performance in the operating room. In: Bogner M, ed. *Human Error in Medicine.* Hillsdale, NJ: LEA 1994;225–53.

8 Healy AN, Undre S, Vincent CA. Developing observational measures of performance in surgical teams. *Qual Safety Healthcare* 2004;**13**(Suppl. 1):i33–40.

9 de Leval MR, Carthey J, Wright DJ, Farewell VT, Reason JT. Human factors and cardiac surgery: a multicentre study. *J Thorac Cardiovasc Surg* 2000;**119**(4):661–72.

10 Joint Committee on Higher Surgical Training (JCHST). Available at: http://www. jchst.org (accessed 22 September 2005).

11 Flin R, Martin L. Behavioural marker systems in aviation. *Int J Aviation Psychol* 2001;**11**:95–118.

12 Flin R, Martin L, Goeters K, Hoermann J, Amalberti R, Valot C, Nijhuis H. The development of the NOTECHS system for evaluating pilots' CRM skills. *Hum Factors Aerospace Safety* 2003;**3**:95–117.

13 Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Anaesthetists' non-technical skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth* 2003;**90**(5):580–8.

14 Rall M, Gaba DM. Patient simulation. In: Miller R, ed. *Miller's Anaesthesia,* 6th edn. Oxford: Elsevier 2005;3073–103.

15 Klampfer B, Flin R, Helmreich RL, Hausler R, Sexton B, Fletcher G, Field P, Staender S, Lauche K, Dieckmann P *et al. Group Interaction in High Risk Environments: Enhancing Performance in High Risk Environments, Recommendations for the Use of Behavioural Markers.* Berlin: GIHRE 2001.

16 Yule S, Flin R, Paterson-Brown S, Maran N. Surgeons' non-technical skills: a review of the literature. *Surgery* 2006;**139**:140–9.

17 Flin R, Yule S, McKenzie L, Paterson-Brown S, Maran N. Attitudes to teamwork and safety in the operating theatre. *The Surgeon* 2006;**4**:145–51.

18 Gordon SE. *Systematic Training Programme Design: Maximising Effectiveness and Minimizing Liability.* Englewood Cliffs, NJ: Prentice Hall 1993.

19 Flanagan J. The critical incident technique. *Psychol Bull* 1954;**51**(4):327–58.

20 Hoffmann R, Crandall B, Shadbolt N. A case study in cognitive task analysis methodology. The Critical Decision Method for the elicitation of expert knowledge. *Hum Factors* 1998;**40**:254–76.

21 Mitiello LG. Representing expertise. In: Salas E, Klein G, eds. *Linking Expertise and Naturalistic Decision Making.* Mahwah, NJ: LEA 2001;245–62.

22 Cooke NJ. Knowledge elicitation. In: Durso FT, ed. *Handbook of Applied Cognition.* New York: Wiley 1999;479–509.

23 Charmaz K. Grounded theory. In: Smith JA, Harre R, Langenhove LV, eds. *Rethinking Methods in Psychology.* London: Sage 1995;26–49.

24 Fletcher G, Flin R, McGeorge P, Glavin R, Maran N, Patey R. Rating non-technical skills: developing a behavioural marker system for use in anaesthesia. *Cogn Technol Work* 2004;**6**:165–71.

25 O'Connor P, Flin R, Fletcher G. Methods used to evaluate the effectiveness of flight crew CRM methods in the UK aviation industry. *Hum Factors Aerospace Safety* 2002;**2**:235–55.

26 Seamster TL, Redding R, Kaempf GL. *Applied Cognitive Task Analysis in Aviation.* Avebury: Aldershot 1997.

27 DeVellis R. *Scale Development: Applications and Theory.* Newbury Park, CA: Sage 1991.

28 Patey R, Flin R, Fletcher G, Maran N, Glavin R. Anaesthetists' non-technical skills (ANTS). In: Henriks K, ed. *Advances in Patient Safety: from Research to Implementation.* Washington, DC: Agency for Quality and Safety in Healthcare 2005;325–336.

29 Catchpole KR, Giddings AE, de Leval MR, Gallivan S, Godden P, Utley M, Dale T, Hirst G. Identifying and reducing systems failures through non-technical skills. Royal College of Surgeons of Edinburgh Quincentenary Congress, June 2005, Edinburgh.

30 Civil Aviation Authority (CAA). *Cap737 – Crew Resource Management (CRM) Training.* Cheltenham: CAA 2003.

APPENDIX 1

Critical incident interview protocol

Social/interpersonal
  Who was in the leadership role? How were they appointed?
  Were any communications required?
  Teamwork within the surgical team
  Teamwork with anaesthetists
  What would have happened if the team members had been different? (i.e. less members, different experience)
  What role did the team have on the case?
  How did you manage resources?
  Was there a briefing?
Cognitive
  What cues were you using to help understand the situation?
  What information did you use in recognizing the situation/making the decision?
  What were your goals during management of the case at this time?
  How did goals alter with the situation?
  How did you re-establish goals?
  What options were open to you at this moment?
  How did you decide which option to take?
  Was there any influence from the team?
  Did you have any heuristics or rules of thumb to use in this situation?
  How did you know they were appropriate?
  How did you arrive at your chosen course of action?
  What factors affected your decision?
  What strategy did you use in reaching your decision?
  Were other team members involved in decision making? (i.e. trainees/anaesthetists)
  On what was your understanding of the unfolding situation based?
  How do you maintain this awareness?
  To what extent did some of your situation awareness come from the team?
  What sort of projections were you making into the future/for what sort of things would you have to be anticipating?
  What are the major elements of which you have to keep track to develop/maintain the big picture?
Analogues/experience
  What could have gone wrong with this operation?
  What would a less experienced surgeon have done in the same situation?
  Do you need different skills to deal with crisis situations as opposed to routine operations?
  What skills do you look for in a good trainee?
  What skills make an excellent surgeon?