

Debriefing surgeons on non-technical skills (NOTSS)

S. Yule · R. Flin · N. Maran · G. Youngson ·
A. Mitchell · D. Rowley · S. Paterson-Brown

Received: 23 February 2007 / Accepted: 29 June 2007 / Published online: 24 August 2007
© Springer-Verlag London Limited 2007

Abstract Surgical trainees must maximise the educational and developmental opportunities of time spent in the operating theatre. Post-operative debriefing on performance based on observed skills is one way of achieving this and is regularly done in other high-risk professions. The non-technical skills for surgeons (NOTSS) behaviour rating system allows surgeons to observe trainees' behaviour in the workplace and provide feedback for skill improvement in a structured manner. This paper describes the process of debriefing using NOTSS and presents the results of a usability trial. Two case studies also illustrate how the system was used. The majority of surgical trainers who participated reported that the NOTSS system provided a common language to discuss non-technical skills and was a valuable adjunct to currently available assessment tools. Some trainers found interpersonal skills more difficult to rate than cognitive skills but 73% ($n = 8$) felt that routine use of the system would enhance patient safety.

Keywords Surgical training · Education · Debriefing · Non-technical skills · Patient safety

1 Introduction

Surgical training in the United Kingdom (UK) has moved away from knowledge-based set pieces in favour of competency-based assessment in the workplace. This is encouraging greater emphasis on identifying the skills necessary to maximise safe and effective management of patients. The competency-based approach requires an appreciation of the role of non-technical skills which are defined as the cognitive and interpersonal skills that underpin technical proficiency (Yule et al. 2006a). There is increasing evidence that the contributory role of non-technical skills failures in surgical adverse events is higher than was ever recognised. Studies conducted in the operating theatre show that that breakdowns in skills such as team working (Schaefer et al. 1995), leadership (Edmondson 2003), communication (Helmreich et al. 2004; Lingard et al. 2004), situation awareness (Way et al. 2003), and decision making (de Leval et al. 2000) are not uncommon and can lead to errors (Stevenson et al. 2007) and poor outcomes (Christian et al. 2006) including avoidable deaths (Templeton and Feinmann 2006), and higher compensation payouts (Studdert et al. 2006). Furthermore, the persistence of these problems may be due to inadequate mechanisms for surgeons to learn from past failures.

Another important development in medical education has been the increasing recognition of the importance of reflection in the training of professionals. In a time of reduced working hours under the European working time directive (EWTD) and exposure to fewer clinical

S. Yule (✉) · R. Flin
School of Psychology, University of Aberdeen,
Aberdeen AB24 2UB, UK
e-mail: s.j.yule@abdn.ac.uk

N. Maran
Department of Anaesthesia, Royal Infirmary, Edinburgh, UK

G. Youngson
Royal Aberdeen Children's Hospital, Aberdeen, UK

A. Mitchell
Aberdeen Royal Infirmary, Aberdeen, UK

D. Rowley
Royal College of Surgeons of Edinburgh, Edinburgh, UK

S. Paterson-Brown
Department of Surgery, Royal Infirmary, Edinburgh, UK

challenges, it is important that clinicians, both in training and career grades, make the most of their clinical experience. Observational research in healthcare is assisting clinicians achieve this by developing tools for activities such as conducting pre-operative briefings (Lingard et al. 2002), structuring observation of behaviours in the operating theatre ‘live’ (Healey et al. 2004) and using video (Guerlain et al. 2004; Mackenzie et al. 2004), as well as debriefing on performance (Yule et al. 2006b). These tools have been designed from a human factors perspective with the aim of improving the safety of surgical patients.

1.1 Briefing in the operating theatre

The main purpose of briefing is to discuss a case before the operation to ensure that the surgical team (surgeons, anaesthetists, nurses, others) are involved, have a shared mental model regarding the goals of the operation, understand specific roles and responsibilities, and are aware of any anticipated complications, technical challenges or unfamiliar equipment to be used. Pre-operative briefings have been developed for use in simulated environments (Gaba et al. 2001) and although not commonly used in surgery, they have been shown to improve interprofessional communication during the operation (Lingard et al. 2005). Briefing is covered in detail in other papers in this special issue so will not be discussed further but one of the important roles of the pre-operative briefing is to set the scene and expectations for the post-operative debrief.

1.2 Debriefing in the OR

Most work on debriefing tends to focus on the aftermath of critical events in military, policing, and health services. Professionals and members of the public may also receive psychological debriefing as part of their treatment for post traumatic stress disorder (PTSD) after traumatic events. This form of debriefing is called event debriefing and is not discussed in the present paper. Instead, this paper focuses on task debriefing which is the process of formally reflecting on performance at work after routine tasks. A formal task debrief involves an experienced professional discussing what happened with the individual(s) who participated in a task in order to help improve their subsequent performance. This is done by identifying aspects of good performance, identifying areas for improvement, and suggesting what should be done differently in future. In order for the experience of debriefing to be meaningful, the person leading the debriefing should have observed the task-related behaviour of others. In many situations, this

person would have been involved in the task in some capacity.

Observation, assessment and structured debrief of observed skills is routine in high risk industries such as European civil aviation (Dismukes and Smith 2000). The most common use of debriefing in healthcare is in anaesthesia, where it is seen as the most important aspect of crisis management training (Rall et al. 2000). Formal task debriefing is not commonplace in surgery but trainees do receive feedback on technical aspects of their surgical skill, for example using the bench model (de Cossart and Fish 2005). Problems encountered during operations may also be discussed in an ad-hoc fashion and mortality and morbidity cases are reviewed in most hospitals. However, routine debrief of performance when operations have been completed successfully does not tend to happen, despite the fact that 73% of trainee surgeons agreed that regular debriefing after an operating theatre session was important for effective team co-ordination (Flin et al. 2006a).

One of the main advantages of debriefing is that it involves expert observation of performance and does not rely solely on self-reflection which can be inherently unreliable (Davis et al. 2006; Gordon 1991). Some surgeons have also been criticised for having a ‘seriously flawed opinion of their own capabilities’ (Templeton and Feinmann 2006), which means that self-reflection alone is unlikely to be an effective method of improving non-technical skills. The large body of research on acquisition and retention of technical skill in surgery stipulates that the process must be an active one with deliberate practice over a sustained period, underpinned by a personal drive for improvement (Guest et al. 2001) and supported by focused feedback and reflection (Regehr and Norman 1996). Feedback on strengths and weakness and self-reflection are likely to be more effective when there is a terminology or vocabulary that permits analysis of performance. One tool which facilitates providing focused feedback on non-technical skills is the non-technical skills for surgeons (NOTSS) system.

1.3 Debriefing on non-technical skills using NOTSS

The NOTSS system was developed using task analysis and allows trained surgeons to structure observations and feedback on observed behaviours in the operating theatre (Yule et al. 2006b). It is based on a skills taxonomy (see Table 1) consisting of four categories (situation awareness, decision making, communication and teamwork, leadership) divided into 12 elements, each with example behaviours. An experimental evaluation found that the NOTSS system had a consistent internal structure and acceptable sensitivity when surgeons’ ratings were

Table 1 NOTSS skills taxonomy v1.2

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
Communication and teamwork	Exchanging information
	Establishing a shared understanding
	Co-ordinating team
Leadership	Setting and maintaining standards
	Supporting others
	Coping with pressure

compared with reference ratings of acceptable/unacceptable behaviour (Flin et al. 2006b). The next stage of usability testing was to determine whether NOTSS could be used by consultant surgeons to observe, rate and debrief trainee surgeons on their non-technical skills after real operations.

Debriefing using NOTSS after a trainee has completed a surgical case allows the trainee to reflect on and critique their performance with the assistance of an expert surgical trainer. During the debrief session, trainees can gain useful insight into achievements, errors and deviations from standards, ask questions and receive specific advice on which to base improvements in performance. Although the focus of the debrief is non-technical skills, medical and technical issues relating to the case are also discussed because the technical and non-technical skills in surgery are intertwined. This meets the criteria for reflection which stipulates that it should be an active process of reflection and be done between cases, following a rigorous process and engaging peers and teachers in dialogue (de Cossart and Fish 2005). One advantage of using NOTSS is that psychometric testing of the system has found high construct validity and a consistent internal structure which means that the NOTSS categories and elements should provide a meaningful structure on which to conduct a task debrief. This paper presents the results of a usability trial and two detailed case studies which illustrate how NOTSS was used to observe and debrief surgical trainees.

2 Method

2.1 Participants

Participants were $n = 22$ surgical trainers (consultant surgeons) and their trainees, recruited from three Scottish

hospitals. Surgical trainers were chosen because it is important that the person leading the debrief session has clinical and teaching experience (Rall et al. 2000). The trainers were asked to use the NOTSS rating form and supporting handbook to rate and provide feedback to trainees as soon as possible after each of ten cases where the trainee had contributed significantly to the operation. Inguinal hernia repair and laparoscopic cholecystectomy were typical operations observed during this trial. Ethical approval to run the study was granted by the Multi-centre Research Ethics Committee (MREC).

2.2 Participant training

Raters must receive specific training to assess non-technical skills (Goldsmith and Johnson 2002). Most consultant surgeons were trained to use the system in a three-hour group session which was based on guidance from aviation on behaviour rating (Baker et al. 2001) and used surgical examples to train raters (Flin et al. 2007). It comprised: (a) background on human factors and non-technical skills; (2) an introduction to the NOTSS system and how to make behavioural assessments, and (3) practice in observing and rating behaviours with NOTSS in three simulated video scenarios. Raters were not calibrated but did discuss their observations and feedback they would give in response to the scenarios. Those who did not participate in this session were given the same training course in a one-to-one setting.

2.3 Usability trial protocol

It was recommended that specific use of NOTSS be determined by the educational needs of the trainees. For example, with junior trainees, the focus of training is on developing basic surgical expertise so it was advised that the NOTSS system be used for general discussion of non-technical skills and their importance to clinical practice. For more senior trainees (i.e. Specialist Registrars), it was suggested that the NOTSS system be used to rate skills and provide feedback during increasingly challenging cases. Surgical trainers were asked to explain to trainees why it is important to assess and provide feedback on non-technical skills during training, highlighting that the NOTSS system has been designed to aid the development of professional skills. Trainees attended an information session about the trial at their hospital. To participate, surgical trainers were asked to:

1. Observe the trainee's behaviours during the operation using NOTSS.
2. At the first suitable point after the operation, rate the trainee's behaviours using a NOTSS rating card and

associated rating scale (4: good, 3: acceptable, 2: marginal, 1: poor, N/A: not applicable).

3. Write feedback notes on the NOTSS rating form.
4. Initiate a discussion with the trainee about the trainee's non-technical performance during the case, using the completed rating form as a basis.
5. Provide the trainee with feedback using a NOTSS form to help him or her improve future performance.

Usability during the trial was assessed using an online post-trial questionnaire of the surgical trainers' views (see <http://www.abdn.ac.uk/iprc/notss> for the questionnaire). An initial invitation to complete it was followed up with a reminder after one month and a further reminder a month later. Self-report measures were selected as the most appropriate method of gathering data on user experiences although are not without limitations—self-report data are by their nature subjective, and susceptible to memory decay and social desirability bias. These are countered in part by an independent observer's account of the debriefing process (see case study 1).

3 Results

3.1 Demographics and implementation of NOTSS

In total, eleven consultant surgeons participated in the usability trial and completed the end-user questionnaire (50% participation rate). Specific data about surgical speciality and gender were not collected as the questionnaire was anonymous but all consultants were in general, orthopaedic or cardiothoracic surgery from three teaching hospitals in Scotland. Data on trainee surgeons were not tracked to ensure that they were confident that the purpose of the study was solely to assess the usability of the tool, rather than their own competence but analysis of completed feedback forms indicate that at least 12 trainees took part.

The NOTSS system was used to observe and debrief on non-technical skills during a total of 43 cases (mean 4 per consultant, range 1–8 cases). In all cases, the trainee was lead surgeon. In some cases the consultant was an unscrubbed observer and on other occasions was scrubbed and assisting as well as observing, but data were not collected on the frequency of this. The majority of trainers preferred to conduct the debriefing immediately after the operation ($n = 9$, 81%) and two trainers (18%) debriefed the trainee at a convenient time later in the same day. Debrief sessions took place in the theatre suite rather than the on the ward or an education/training room. Five trainers (45%) conducted the debriefing in the operating theatre itself; two (18%) used an office in the theatre suite, two (18%) debriefed the trainee in the coffee room, and one (9%) used the anaesthetic room. One trainer did not respond to this question. The median length of debrief session was 3–5 min (six trainers, 54%), two trainers averaged debrief sessions less than 2 min (18%) and two averaged 6–10 min (18%). One trainer's average debrief time was over 16 min (9%).

See Fig. 1 for an example of a completed NOTSS rating card. The feedback notes were used as a basis to conduct a debrief session after an inguinal hernia repair which mainly focused on the trainee's ability to gather information, consider options regarding the optimal suture and mesh size for the patient, and ability to liaise with the anaesthetist about the patient's condition. This rating card was completed during one of the feedback sessions described in case study 1.

Ten trainers (90%) thought that they had received just enough training to use the system and one trainer (9%) thought that he/she had not received enough training. The majority of trainers used both categories and elements to structure observation and debrief (nine trainers, 81%) and two trainers (18%) focused on the categories only. None of the trainers reported focusing on just the elements. Of individual categories, all trainers used 'communication and

Fig. 1 Completed NOTSS rating form for trainee performance on right inguinal hernia (RIH) repair—see also case study 1

Hospital: Date:		Trainer name: Trainee name:		Operation: Right inguinal hernia repair
Category	Category rating*	Element	Element rating*	Feedback on performance and debriefing notes
Situation Awareness	3	Gathering information	3	Didn't mark side/ arrived in theatre late
		Understanding information	4	Aware of INR importance and checked
		Projecting and anticipating future state	3	Take more of a lead in op - i.e. requesting retractions
Decision Making	3	Considering options	3	Be more explicit about relative merits of options
		Selecting and communicating option	2	Not sure about sutures/ mesh sizes etc...
		Implementing and reviewing decisions	4	Readily vocalised concerns
Communication and Teamwork	2	Exchanging information	2	Did not relate well to anaesthetist
		Establishing a shared understanding	3	Waited for trainer to take the lead
		Co-ordinating team activities	3	Did not enquire about pt condition from anaesthetist
Leadership	3	Setting and maintaining standards	4	Followed theatre protocol but didn't mark side
		Supporting others	4	Good rapport with trainee OPD and scrub nurse
		Coping with pressure	3	At times seemed to carry on regardless - oblivious to important anatomy - too focused on other things

teamwork', 10 (90%) used 'situation awareness', eight (72%) used decision making, and just over half ($n = 6$, 54%) used the leadership category. Some categories were not used by some trainers due to the level of the trainee and the complexity of the procedure being completed. For example, one trainer stated that the cases selected for his/her trainee were not amenable to a good discussion regarding decision making.

3.2 System usability

The results of initial usability questions are presented in Table 2. The majority of surgical trainers thought that the NOTSS system was useful for debriefing trainees and a valuable adjunct to currently available assessment tools. The trainers were all in agreement that NOTSS provided a common language to discuss non-technical skills and was useful to support reflective practice. Implementation of NOTSS was positive but with mixed opinions regarding the ease of rating skills. Although 45% of trainers agreed that cognitive and interpersonal skills were easy to rate, 27% found interpersonal skills difficult to rate compared with only 9% who felt cognitive skills were difficult to rate. The remaining trainers were ambivalent regarding ease of rating. Time can be a precious commodity in the operating theatre but only 9% of trainers thought using NOTSS to debrief added too much time to their operating list. And although 18% of trainers disagreed, 73% thought that

routine use of NOTSS would enhance safety in the operating theatre. All trainers thought that NOTSS has a place in surgical education and assessment. Comments from trainers indicated that positive aspects of the system for surgical education were the transparent structure; common language; ability to objectively assess skills; framework for providing feedback; ease of use in real life situations, and that using the system made time to discuss aspects of surgical performance that are 'usually ignored'.

Although some trainers reported no difficulties rating behaviours using NOTSS, four main difficulties were articulated. These related to understanding some descriptors in the NOTSS handbook; selecting an appropriate trainee and case; observing and rating behaviours while also scrubbed, and an over-reliance on communication to infer cognitive skills. Consultant surgeons also reported that they used NOTSS to discuss problems that trainees encountered and to highlight the importance of planning procedures in advance. In-depth accounts of how NOTSS was used can be found in the case studies in the following section.

3.3 Case studies documenting how NOTSS has been used

To illustrate how the NOTSS system has been used in the clinical environment, two case studies were written. The first documents how NOTSS was used to provide regular

Table 2 Initial usability results ($n = 11$ consultant surgeons)

Question	Strongly disagree (%)	Disagree (%)	Neither (%)	Agree (%)	Strongly agree (%)	N/A
The NOTSS system was a useful tool for debriefing trainees after the operation	–	–	–	73	27	–
The NOTSS system provided a common language to discuss non-technical skills	–	–	–	54	36	9
It was easy to rate cognitive skills (situation awareness, decision making)	–	9	45	45	–	–
It was easy to rate interpersonal skills (communication and teamwork, leadership)	9	18	27	36	9	–
Using NOTSS added too much time to my operating list	27	45	18	9	–	–
NOTSS is a useful tool to support reflective practice	–	–	–	73	27	–
The NOTSS system is a valuable adjunct to the available assessment tools	–	–	9	54	36	–
Routine use of the NOTSS system will enhance safety in the operating theatre	–	18	9	64	9	–

feedback after routine operations and is written from the perspective of the researcher who observed the process. Although it describes only one case, it illustrates the benefits of receiving regular debriefing on non-technical skills even when things appear to be going well. The second case study was written by a surgical trainee who documented his experience of using NOTSS with a trainer to analyse and learn from an adverse event (injury to wall of colon). The NOTSS system was not specifically developed to investigate failures, but as this emergent use of the system was instigated by surgeons it was important to document when addressing initial system usability.

3.3.1 Case study 1: Using NOTSS to conduct a regular post-operative debrief

Patient: Four male patients

Operation: Right inguinal hernia repair (×4)

Trainee level: Specialist Registrar (SpR)

Case study written by a researcher who observed the process

Before the case A General surgery consultant ('the trainer'), trained to use NOTSS, selected a senior trainee (Specialist Registrar) to trial the NOTSS system with him. The trainee was technically competent and confident, and had not been selected because his non-technical skills were perceived to be substandard. Trainer and trainee both reviewed the NOTSS system handbook and rating/feedback form. It was decided to trial NOTSS during and after a right inguinal hernia repair.

On the day of the case After quickly discussing activity on the surgical ward, attention turned to the operative case. The trainer checked that the trainee had read the NOTSS handbook and was familiar with the terminology. The trainer then confirmed that the trainee would be doing the operation and that he (the trainer) would be an unscrubbed observer. The trainer described how the debrief would proceed: it was agreed that the trainee would start by describing his performance using the NOTSS categories and elements, then the trainer would add his observations and ratings, and finally provide feedback for improvement. The trainer and trainee then discussed technical aspects of the hernia repair and decided which mesh size they would likely use. They then discussed the rest of the operating list, deciding to re-order the operations because one patient was diabetic but listed towards the end of the day. They reconvened in the operating theatre once the patient was anaesthetised.

During the operation The trainer observed the trainee's behaviour and referred to the NOTSS handbook as required. He did not take notes during the case. Although

unscrubbed, the trainer did occasionally advise the trainee on technical aspects of the operation.

Immediately post-operatively The trainer and trainee ensured that the patient reached recovery and then dictated the operation notes. They then went to a small private office adjacent to the operating theatre for the debrief, which took place about 10 minutes after the case had been completed. The trainee began, and focused on the NOTSS categories to describe his non-technical skills at key memorable (technical) phases of the operation. The trainer then added his perspective, rated the trainees' behaviour using the 1–4 scale, and justified the ratings based on his observations. The trainee responded to some of the feedback (e.g. regarding choice of sutures and mesh) and was sometimes defensive but the trainer seemed to accept those justifications and there was an open discussion about the case. The trainer then provided expert feedback and guidance for improvement which was recorded on a NOTSS feedback card and kept by the trainee for future reference (see Fig. 1). The debrief lasted 12 minutes.

During the week after the case A further three cases were observed in next 10 days, following the same process. During the feedback sessions the trainer and trainee both made reference to previous cases to illustrate similarities and differences in the trainee's behaviour. The rating cards and feedback notes from previous cases were at hand and occasionally consulted.

3.3.2 Case 2: Using NOTSS to debrief after a surgical adverse event

Patient: 12 year old male

Operation: laparoscopic appendicectomy

Trainee level: Specialist Registrar (SpR)

Case study written by the trainee surgeon

I conducted a structured analysis and review of an operative complication using the NOTSS framework. The procedure was a laparoscopic appendicectomy which I undertook without the presence of the patient's consultant but with another consultant surgeon observing in theatre. The appendix was found behind the colon in a higher than normal location increasing the complexity of the procedure. After reaching a stage where I felt I could not continue, I dissected further under my senior colleague's instruction but still did not make progress. My colleague then scrubbed and dissected during which an iatrogenic injury to the wall of the colon occurred. There was intra-operative recognition of this complication and appropriate management following conversion to an open procedure. The patient made an unremarkable recovery and was discharged without significant delay.

NOTSS-based feedback was conducted with my trainer (the patient's consultant) several days following the event. Mixed experience of feedback situations made me adopt a defensive approach. My trainer selected a quiet neutral venue rather than intimidating surroundings of his office for the session. We used the NOTSS framework to facilitate a qualitative, formative and constructive analysis of the operative complication. The discussion identified that situation awareness during the case was satisfactory—it was recognised before and during the procedure that this would not be an easy appendectomy for technical reasons. Situation awareness remained satisfactory following the complication, with projection of the future state and appropriate action taken. Regarding decision making, we discussed how I considered conversion to an open procedure when I was no longer making progress with the dissection. I recognised that completion of the procedure laparoscopically would be in the patient's interest and decided to transfer the role of operating surgeon to a senior colleague who offered to continue the procedure laparoscopically. We concluded that appropriate decisions were made but that the timing of decision making was adversely affected by communication and leadership failures. There was a lack of information exchange at the time when I felt we should have converted to an open procedure. We did not have a shared understanding of the apparent danger of continued dissection without more accurate appreciation of the anatomical planes and I felt unable to challenge the senior colleague's decision making strongly enough. This related directly to the NOTSS element of setting and maintaining standards. Our analysis of communication and teamwork and leadership also raised satisfactory aspects such as good coordination of team activities during the process of conversion to the open procedure, rectification of the injury, and completion of the procedure.

The feedback process reached far deeper in the analysis of situation, causation and process of the adverse event than the superficial technical evaluation I had anticipated. The outcome enabled me to extract positive constructive learning value from the process with wider application to my general operative approach and I feel that it enhanced the development of my non-technical skills.

4 Discussion

The surgical trainers who used NOTSS to observe, rate and debrief their trainees in this trial were very positive about the usability of the system and thought that NOTSS had a place in surgical education. The majority of trainers used the 'communication and teamwork' and 'situation awareness' categories but there was less agreement about whether 'leadership' and 'decision making' were suitable

for trainees. For leadership, the mere presence of the consultant in theatre may have the effect of marginalising the trainee's ability as leader, even if the trainee is lead surgeon in the operation. Deference to the consultant surgeon may have the effect of inhibiting the trainees' interpersonal skills because they assume that the consultant is in the leadership role and will therefore communicate critical information with other team members. Surgeons who are scrubbed also seem more likely to get involved in the operation too readily which has the effect of reinforcing the power imbalance between trainer and trainee. The differences in use of the decision making category between consultants may reflect differences in case-mix. The operations selected by consultants for this trial such as laparoscopic cholecystectomy perhaps required less intra-operative decision making than more complex cases, so the decision making category appeared to be not required or was underutilised.

Four main difficulties using the system were articulated. Firstly, difficulties understanding the descriptors in the NOTSS handbook were reported. It is likely that those issues would be addressed with more training but will also be addressed when the prototype tool is revised. Secondly, trainers sometimes found it difficult to select an appropriate trainee to trial the system with. Trainers had not been given specific guidance about this and a high level of technical competence is required before professionals' non-technical skills can be rated in the workplace or simulated environment. Therefore, the system appears more suitable for use with higher surgical trainees (i.e. Specialist Registrars) because they are less concerned with learning technical skills than more junior trainees are. Choice of an appropriate case or theatre list also proved difficult in some instances. Because the NOTSS system is only valid for observing surgeons while they are operating, the case chosen had to be appropriate for their level of competence. Allied to this is role of the consultant in theatre. Ideally they would be an unscrubbed observer in the operating theatre but on several occasions, the trainers were scrubbed and assisting the trainee, which can have a negative impact on their ability to observe and rate behaviours and may also interfere with the task and trainee's ability to assume a leadership role in theatre. The importance of communication and verbalisation was highlighted by one trainer who noted difficulties rating non-technical skills when the trainee did not verbalise adequately. Using the NOTSS system does rely heavily on observing communication behaviours in order to infer cognitive skills (situation awareness and decision making) but not exclusively so. For example, non-verbal behavioural markers of situation awareness include 'looking at patient notes and imaging', and 'selecting inappropriate manoeuvre' and 'becoming rushed' are indicators of decision making. Specific training in

identifying these non-verbal behaviours would help observers rate skills more confidently in the absence of communication.

The majority of surgeons thought that routine use of the NOTSS system would improve patient safety. However, one potential reason why the system may not have universal appeal is that it will not in itself make healthcare safer as using NOTSS requires longer-term involvement and effort from surgeons. The effects of regular debriefing are not only long term, as there are short-term learning experiences that can be identified one day and put into practice in the next case. However, there is an effort–reward balance, and some surgeons may favour other interventions to improve safety, or believe that patient safety is at an acceptable level, therefore, there is no need for this type of research.

4.1 Study limitations

Despite initial enthusiasm, 50% of the trainers did not fully participate in the trial and complete a usability questionnaire, and of those who did participate, none achieved the target of using the system during ten cases. Although data from those trainers were not available, aspects of system usability or implementation may have contributed to the 50% attrition rate.

4.1.1 Design of the trial

There are several possible reasons for this level of non-participation: (a) it was hard to make time for educational activities that focus on the process of surgery, rather than the technical aspects of operating, (2) surgeons did not want to routinely debrief or be debriefed about their work, or (3) it was difficult to persuade surgeons to adopt a new system. It is most likely a combination of all three, underpinned by cultural resistance to this type of activity (i.e. debriefing on non-technical skills is not ‘the way we do things round here’) despite enthusiasm at the individual level. The types of cultural barriers encountered by this study were the autonomy of trainers (high-level support from senior surgeons was necessary but not sufficient and participation had to be negotiated on an individual basis); the novelty of the system (the format of NOTSS was not analogous to the technical training or experiences of trainers), and that the system was used when time permitted, rather than having a dedicated slot in particular operating lists (thus was used rarely as other matters relating to patient care took precedence). Surgery is a dynamic process and the NOTSS handbooks were not always available when the opportunity arose for a

debriefing (i.e. the surgeon’s personal handbook was often in the office or another theatre), resulting in a lost opportunity. As non-technical skills language becomes more common in surgery, and part of ‘the way things are done round here’, the cultural barriers to participation in this type of activity will dissipate. It should also be noted that healthcare and surgery are not unique in this respect—pilots do not observe, rate and provide feedback instinctively but have to be persuaded to engage in this process even if it is mandated in some countries (Dismukes and Smith 2000). Usability does not exist independently from context and the results from those who completed the trial are promising. However, the sample size is not sufficient and more data needs to be collected in order to draw definitive conclusions about the usability of the system, especially as those who did participate were prior advocates of the system. Some lessons have been learned. To increase participation in the use of NOTSS, an on-site facilitator is required to arrange debrief sessions and provide materials during the trial in addition to providing the materials and training at the start of the trial. More NOTSS materials and publicity should also be available in the operating theatres, rather than relying on trainers to carry their handbook and rating forms with them at all times.

Alternative raters could also be considered. NOTSS was designed to be used by surgeons with technical knowledge of surgery but with appropriate training, people who are familiar with the operating theatre environment and an understanding of the phases of intraoperative surgery could credibly observe and rate behaviours. In this respect, professionals such as anaesthetists, scrub nurses, behavioural scientists or communication experts may be suitable, but this would have to be tested empirically. Surgeons and trainees have already acknowledged a willingness to receive feedback from individuals out with their profession (Flin et al. 2006a). Future studies must be more cognizant of cultural aspects of the working environment in order to be successfully implemented.

4.1.2 Lack of training for trainers

Some trainers reported finding NOTSS difficult to pick up and use. It should be noted that although the system appears deceptively simple, specific training must be undertaken by potential system users. Trainers had attended a two and a half an hour training session on non-technical skills but despite this being the maximum amount of training that it was possible to give consultant surgeons, it is less than that recommended for training professionals to use behaviour rating systems in other domains (Baker et al. 2001; Goldsmith and Johnson 2002; Klampfer et al. 2001). In fact, it is suggested that a minimum of two

consecutive days training is received to use the Anaesthetists' non-technical skills (ANTS) system (Patey et al. 2005), and debriefing skills should also be explicitly trained (Rall and Gaba 2005). It was assumed (wrongly in some cases) that experienced consultant surgeons would be competent at providing feedback without training. Non-technical language is not yet common parlance in surgery and we assumed that surgeons would automatically adopt this 'new' language without practice. Ability to use appropriate language for feedback and setting the scene and trainee expectations for feedback plays a large part in whether the debrief session is successful, and is as important as the ability to observe and rate the relevant behaviours. This does not mean that feedback skills are generic—it is necessary to have a language for feedback (i.e. NOTSS) before training in how to provide feedback can be given. It is suggested that future trials train system users to provide feedback using the specific language and behavioural markers in the NOTSS system.

5 Conclusion

Usability of the NOTSS system was very positive but only from a small sample of committed consultants so this may not be reflective of the wider view in surgery. It is perhaps unrealistic for a new system like NOTSS to be adopted in surgery unless it is taken seriously and use of the system to debrief or assess non-technical skills mandated in some form. With increasing experience of using the system, surgeons and trainees will become more familiar with the terminology and more confident in using NOTSS to observe, rate and provide feedback. The case studies illustrate the potential use and benefits of a non-technical skills framework in this respect. Current proposals to overhaul medical regulation in the UK (Donaldson 2006) means that non-technical skill competencies could be the subject of formal assessment in the future. Surgery can learn from other high reliability industries whom routinely brief and debrief personnel as a way to enhance safety. If non-technical skills are to be formally addressed in surgical education, then routine use of NOTSS will be able to provide trainees with more regular feedback on their non-technical skills and trainers will have a more solid and defensible evidence-base on which to make formative and summative assessments.

Acknowledgments The NOTSS system was developed under funding from the Royal College of Surgeons of Edinburgh and NHS Education for Scotland. The views presented in this manuscript are those of the authors and should not be taken to represent the position or policy of the funding bodies. We would like to thank the surgeons who took part in this study.

References

- Baker D, Mulqueen C, Dismukes R (2001) Training raters to assess resource management skills. In: Salas E, Bowers C, Edens E (eds) *Improving teamwork in organizations*. LEA, New Jersey, pp 131–145
- Christian C, Gustafon M, Roth E, Sheridan T (2006) A prospective study of patient safety in the operating room. *Surgery* 139:159–173
- Davis DA, Mazamian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L (2006) Accuracy of physician self-assessment compared with observed measures of competence. *J Am Med Assoc* 296:1094–1102
- de Cossart L, Fish D (2005) *Cultivating a thinking surgeon*. tfm publishing, Shrewsbury
- de Leval MR, Carthey J, Wright DJ, Farewell VT, Reason JT (2000) Human factors and cardiac surgery: a multicentre study. *J Thorac Cardiovasc Surg* 119(4):661–672
- Dismukes R, Smith GM (2000) *Facilitation and debriefing in aviation training and operations*. Ashgate, Aldershot
- Donaldson LJ (2006) *Good doctors, safer patients: proposals to strengthen the system to assure and improve the performance of doctors and to protect the safety of patients*. Department of Health, London
- Edmondson AC (2003) Speaking up in the operating room: how team leaders promote learning in interdisciplinary action teams. *J Manage Stud* 40(6):1419–1452
- Flin R, Yule S, McKenzie L, Paterson-Brown S, Maran N (2006a) Attitudes to teamwork and safety in the operating theatre. *Surgeon* 4:145–151
- Flin R, Yule S, Paterson-Brown S, Maran N, Rowley D, Youngson GG (2006b) Experimental evaluation of a behavioural marker system for surgeons' non-technical skills (NOTSS). Paper presented at the 50th annual human factors and ergonomics society meeting, San Francisco
- Flin R, Yule S, Paterson-Brown S, Maran N, Rowley DR, Youngson GG (2007) Teaching surgeons about non-technical skills. *Surgeon* 5(2):86–89
- Gaba DM, Howard SK, Fish KJ, Smith BE, Sowb Y (2001) Simulation-based training in anaesthesia crisis resource management (ACRM): A decade of experience. *Simul Gaming* 32:175–193
- Goldsmith T, Johnson P (2002) Assessing and improving evaluation of aircrew performance. *Int J Aviat Psychol* 12:223–240
- Gordon MJ (1991) A review of the validity and accuracy of self-assessments in health professions training. *Acad Med* 66(12):762–769
- Guerlain S, Turrentine B, Adams R, Calland JF (2004) Using video data for the analysis and training of medical personnel. *Cognit Technol Work* 6:131–138
- Guest C, Regehr G, Tiberius R (2001) The life long challenge of expertise. *Med Educ* 35:78–81
- Healey AN, Undre S, Vincent C (2004) Developing observational measures of performance in surgical teams. *Qual Saf Health Care* 13(Suppl 1):i33–i40
- Helmreich R, Musson D, Sexton B (2004) Human factors and safety in surgery. In: Nora P, Manuel B (eds) *Surgical patient safety: essential information for surgeons in today's environment*. American College of Surgeons, Chicago
- Klampfer B, Flin R, Helmreich RL, Hausler R, Sexton B, Fletcher G et al (2001) Group interaction in high risk environments: enhancing performance in high risk environments, recommendations for the use of behavioural markers. GIHRE, Berlin
- Lingard L, Reznick R, Devito I, Espin S (2002) Forming professional identities on the health care team: discursive constructions of the 'other' in the operating room. *Med Educ* 36:728–734

- Lingard L, Espin S, Whyte S, Regehr G, Baker G, Reznick R et al (2004) Communication failures in the operating room: an observational classification of recurrent types and effects. *Qual Saf Health Care* 13:330–334
- Lingard L, Espin S, Rubin B, Whyte S, Colmenares M, Baker GR (2005) Getting teams to talk: development and pilot implementation of a checklist to promote interprofessional communication in the OR. *Qual Saf Health Care* 14:340–346
- Mackenzie CF, Xiao Y, Horst R (2004) Video task analysis in high performance teams. *Cogn Technol Work* 6:139–147
- Patey R, Flin R, Fletcher G, Maran N, Glavin R (2005) Developing a taxonomy of anaesthetists' nontechnical skills (ANTS). In: Hermison K (ed) *Advances in patient safety: from research to implementation*. AHRQ, Washington
- Rall M, Gaba D (2005) Human performance and patient safety. In: Miller R (ed) *Miller's anesthesia*. Elsevier, Philadelphia
- Rall M, Manser T, Howard SK (2000) Key elements of debriefing for simulator training. *Eur J Anaesthesiol* 17:516–517
- Regehr G, Norman GR (1996) Issues in cognitive psychology: implications for professional education. *Acad Med* 71:988–1001
- Schaefer H, Helmreich R, Scheidegger D (1995) Safety in the operating theatre—Part 1: Interpersonal relationships and team performance. *Curr Anaesth Crit Care* 6:48–53
- Stevenson KS, Gibson SC, MacDonald J, Hole DJ, Rogers PN, Byrne DS et al (2007) Measurement of process as quality control in the management of acute surgical emergencies. *Br J Surg* 94(3):376–381
- Studdert DM, Mello MM, Gawande AA, Gandhi TK, Kachalia A, Yoon C et al (2006) Claims, errors, and compensation payments in medical malpractice litigation. *N Engl J Med* 354(19):2024–2033
- Templeton S, Feinmann J (2006). Arrogant surgeons 'risk another Bristol babies scandal'. *The Sunday Times* (3 September 2006)
- Way L, Stewart L, Gantert W, Kingsway L, Lee C, Whang K et al (2003) Causes and prevention of laparoscopic bile duct injuries. *Ann Surg* 237:460–469
- Yule S, Flin R, Paterson-Brown S, Maran N (2006a) Non-technical skills for surgeons: a review of the literature. *Surgery* 139:140–149
- Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D (2006b) Development of a rating system for surgeons' non-technical skills. *Med Educ* 40:1098–1104