Ergonomics
Vol. 49, Nos. 5–6, 15 April–15 May 2006, 589–604

Measuring intra-operative interference from distraction and interruption observed in the operating theatre

A. N. HEALEY*, N. SEVDALIS and C. A. VINCENT

Clinical Safety Research, Imperial College, University of London, Department of Biosurgery and Technology 10th Floor QEQM, St Mary's Hospital, Praed Street, Paddington, London W2 1NY, UK

An observational tool was developed to record distraction and interruption in the operating theatre during surgery. Observed events were assigned to pre-defined categories and rated in relation to the level of team involvement – the sum of which was treated as a measure of intra-operative interference. Many events (0.29 ± 0.02 per min) were observed and rated in 50 general operations sampled from a single operating theatre. The rating of individual events ($r_s = 0.65$) and of cases ($r_s = 0.89$) correlated between independent observers. Interference levels (1.04 ± 0.07/min) also correlated with door opening frequency (0.68 ± 0.03/min) ($r = 0.47$, $p < 0.001$). Some sources of interference were intrinsic to the work of the surgical team, including equipment, procedure and environment, while others were extraneous, including bleepers, phone calls and external staff. The findings highlight the need to further develop measures of interference, to assess its variation, intensity and its effect on surgical team performance.

Keywords: Operating theatre; Surgery team; Distraction; Interruption; Performance; Safety

1. Introduction

A systems view is increasingly important to health care, for improving efficiency and patient safety (Department of Health 2000). Safety from the systems perspective is achieved by improving the interaction among system components (Reason 2000, Wilson 2000). There is considerable untapped potential for studying system interactions within the operating theatre in order to highlight weaknesses in the process of care. This study was an attempt at developing a measure of intra-operative interference from distraction and interruption observed in the operating theatre. The study was motivated by the need to improve team performance and reduce adverse events in surgery (Brennan *et al.* 1991).

*Corresponding author. Email: a.healey@ic.ac.uk

Ergonomics
ISSN 0014-0139 print/ISSN 1366-5847 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00140130600568899

Multidisciplinary teamwork in surgery poses a challenge to those aiming to optimize the performance of surgical teams. Teamwork in surgery involves collaboration among anaesthetists, nurses, surgeons and associated personnel. Team members are jointly responsible for patient care, yet at the same time they have different multiple tasks to perform in shared space in the operating theatre. The daily caseload of surgical teams is such that teamwork activity between surgical cases may encroach upon activity within cases; different tasks and activities in a team may interfere with one another. In other words, the conditions necessary for the successful execution of one task may be detrimental to those of another task and there may be numerous forms of distraction and interruption in the operating theatre.

Researchers have studied the effects of multiple tasks in work environments, where certain tasks and activities distract a person from a primary task or interrupt their task momentarily. For instance, distraction and interruption has been measured in the office environment (Quintus and George 2003, Banbury and Berry 2005) and in the aviation cockpit (Latorella 1996, Dismukes *et al.* 2001) using observational and self-report methods. These studies show that distraction and interruption usually have some detrimental effect on work. In the cockpit, distractions and interruptions can negatively affect a pilot's performance in completing safety checks (Latorella 1996) and disrupt their prospective memory; in that they forget their intentions for planned events (Loukopoulos *et al.* 2001, Einstein *et al.* 2003). In some circumstances, distraction or interruption may contribute to aviation accidents (Chou and Funk 1990). Behavioural adaptations may counteract the detrimental effects of distraction or interruption in the short term (Dismukes *et al.* 2001), but adaptation itself may degrade performance in the long term (Zijlstra *et al.* 1999).

Research has also shown that aspects of the physical and social environment that can interfere with the work of operating theatre teams (Weinger and Englund 1990, Sexton *et al.* 2000). New technology, such as that which affords minimal-access surgery, introduces considerable ergonomic problems for a range of disciplines in the operating theatre (Welty *et al.* 2002, Van Veelen *et al.* 2003, Gerbrands *et al.* 2004). However, there are few studies that have focused primarily on assessing interactions in surgery and very few that have focused on distraction or interruption in the operating theatre.

In a relevant study of an emergency department, interruptions and breaks in primary task activity experienced by physicians during 180-min periods of work were recorded by observation (Chisholm *et al.* 2000). Interruptions were defined as any event demanding a physician's momentary attention. Breaks were defined as interruptions lasting more than 10 s that resulted in a task switch. The observers recorded a mean of 30.9 interruptions and 20.7 breaks during 180-min periods. When combined, this approximates to a rate of 0.29 interrupting events per min. The authors concluded that interruptions were necessary to meet the demands of multiple cases but that excessive interruption may hinder a physician's performance.

This study set out to observe and record the frequency of distraction and interruption in the operating theatre during the intra-operative phase of surgery, that is, from incision to closure (pre- and post-operative phases warrant separate study). Distraction was defined as a break in attention, evidenced by observed behaviour, such as orienting away from a task or verbal responding. Interruption was defined as a break in task activity, evidenced by observed cessation of a task. It was helpful, indeed important, to view teamwork in theatre as a system confined to single surgical cases in order to disambiguate work from distraction and categorize events accordingly. The team was defined as the personnel assigned to a surgical case, comprising three main groups: anaesthetists; surgeons; nurses and their respective assistants.

Of course, some types of distraction and interruption are likely to interfere with the team's work more than other types. Some types might consistently involve just one team member, while others might involve the whole team. Therefore, observed events were weighted using a rating scale that related to how many team members were involved in an event. For each case, the sum of all rated events formed a measure of intra-operative interference. The counts and ratings enabled a comparison across cases and an analysis of the relationship of interference to other variables, such as disciplinary group, operation type, the rate of door opening and ambient noise level in the operating theatre.

2. Method

2.1. *Sample*

A total of 50 general operations (29 laparoscopic and 21 open) were sampled from a single operating theatre in a National Health Service Teaching Hospital. Ethics approval was gained from the appropriate panel. Patient consent was not sought as patient details and identities were not gathered. The observers were present in the capacity of researchers employed within a surgical department. Short examinations and cases likely to last more than 4 h, from incision to closure, were excluded, as the observational method was particularly demanding of attention. Data were collected on Tuesday, Thursday and Friday of each week on available operations over a 3-month period. Team composition varied: different surgeons were allocated to those 3 d and anaesthetists tended to vary from case to case. However, there was some consistency in nursing personnel, as some nurses were assigned to the operating theatre sampled. Two research psychologists shared responsibility for data collection. They collected data jointly on 11 operations, four in each of the first 2 months and three in the third month, in order to run tests of inter-rater reliability on distraction levels for individual events and for ordinal discrimination of sum levels per case.

2.2. *Procedure*

Previous experience in carrying out research in this area (Healey *et al.* 2004) guided the initial construction of a data record sheet with predefined categories and an ordinal rating scale to assign interference ratings (table 1). Observer 1 had previously observed >50 surgical operations in another study while Observer 2 had observed six surgical cases. Observer 1 piloted the observational measure over three sessions in theatre. In a further six sessions, observer 1 and observer 2 piloted the measure jointly to gain a shared understanding of recording events and assigning ratings to them. Each distracting or interrupting event was rated according to the number of personnel involved in the event as shown in table 1.

Scale points 1–3 refer to salient events that potentially or actually distract or interrupt the work of a circulating nurse. The first scale point related to potential distraction. That rating was assigned to observed events that may be distracting, but where no behavioural evidence of distraction was observed. For example, it is possible that a bleeper may activate, but no one responds to it; this would be rated as 1. A bleeper might, however, be noticed by a floating team member but not dealt with, which would rate 2 on the scale. If a circulating nurse dealt with the bleeper, that would score 3 on the scale.

Scale points 4–6 refer to observed distraction or interruption to a single member of the team. Specifically, point 4 refers to an observed distraction evidenced by observed

Table 1. Ordinal scale used to rate observed sources of interference*.

Level	Observed effects on team
1	Potentially distracting source
2	Interference noticed by floating personnel
3	Floating member attends to non-case interference
4	Team member momentarily distracted from task
5	Team member pauses current task
6	Team member attends to distraction
7	Team distracted momentarily
8	Team attend to distraction
9	Operation flow interrupted

*The scale points relate to the team's involvement in an event. Lower scores 1–3 are related to salient stimuli or particular instances that may affect the team, such as bleepers that are either ignored or dealt with by circulating personnel, respectively. Scale points 4–6 relate to individual members being distracted by an event. Highest scale points 7–9 relate to two or more members of the surgical team being distracted, leading to workflow interruption.

behaviour, such as orienting away from a task, perhaps to a sound or a visually detected movement, while continuing with a task. Point 5 refers to distraction where a team member is observed to pause his or her task for a moment, for example, where the surgeon conducting laparoscopy turns to address the instrument tray to view the equipment available, while retaining control of instruments inserted in the patient's abdomen. Point 6 refers to an event that interrupts a team member's work, for example, where the anaesthetist is called to attend to queries about another case.

Scale points 7–8 refer to similar distraction or interruption to points 5 and 6, but where two or more team members are involved. The highest scale point 9 refers to observed interruption to the whole team, where they are observed to attend to another event. For example, an external member of staff may enter to discuss another matter with several members of the team or a problem in theatre may be detected, perhaps an unusual sound or failure in equipment.

It was intended that any salient source of interference observed would be recorded but to aid data recording, it was necessary that observers use a pre-specified list of categories that events might refer to, namely:

1. Phone – any phone in theatre or next to theatre.
2. Bleeper – any bleeper activated in theatre.
3. Radio – action or response to the radio causing distraction.
4. Case irrelevant communication – any conversation irrelevant to the case.
5. Communication difficulties – e.g. lack of response to request.
6. External staff – anyone not part of the team in theatre (except the observer).
7. Equipment – any item of equipment or provision not at hand or failing.
8. Work environment – workspace and human–interface problems.
9. Procedural – events intrinsic to the case work.
10. Movement in front of or behind video monitors – laparoscopic cases only.

Of course, it is evident that a single distracting event may be placed into more than one category; there is inevitable overlap. However, the aim was to consider the initiating source of interference. In other words, the initial prompt of a distracting event was used to categorize events. For example, external staff may call from an operating theatre door and begin a conversation with the surgeon. In that case, the conversation may be distracting, but external staff at the door was the initial stimulus. If during the operation an item of equipment caused interruption that event would be recorded in the equipment category. However, the equipment failure might have been due to a lack of pre-operative preparation or an ineffective maintenance check. Both of those causes might refer to a failure in procedure. The point is that the criterion for categorization was based upon 'observations' not 'inference'.

Data collection was limited to the intra-operative phase only, that is, from incision to closure. Data collection was therefore intentionally focused on the team's activity during the course of a surgical operation. For each case, the time of incision and time of closure was noted. Observed events were matched to the list and rated using the 9-point scale. The time of the event was recorded to aid cross-referencing of distractions for inter-observation reliability testing. Those involved and affected by the event were also recorded. Three groups were defined in the study. The anaesthetist group comprised anaesthetists and their assistants. The nursing group comprised sterile nurses, circulating nurses and any assistants. The surgical group comprised surgeons operating and assisting and any surgeons or surgical trainees accompanying the former in the operating theatre.

Sound pressure levels (dB(A)) were recorded at 2-s intervals with a Tecpel SE-322 sound level meter (Tecpel Co. Ltd, Taipei, Taiwan), auto ranging measurement between 30dB(A) and 130dB(A), accuracy $\pm$ 1.5dB(A) and a capacity for 32 000 data logs. Finally, a tally of staff flow in and out of theatre in terms of door-opening frequency was also recorded. This additional measure was taken in parallel to the recording of observed events.

The observers positioned themselves in theatre so that all team members could be observed (see figure 1); observer position once established was usually static. A primary objective for observers was not to interfere with the work of operating theatre personnel. In the first instance, the theatre manager was asked where the observer could be positioned. The observers experienced very few enquiries into the nature of their work from personnel; the personnel seemed habituated to observation, perhaps from regular visits by students and clinicians. Whenever a member of the operating theatre team inquired about the observer's reasons for being in theatre, a very brief explanation was provided.

2.2.1. Data tabulation. Data were tabulated during observations on standard data record forms containing separate columns for the following:

1. Source category (ten categories described in 2.2).
2. Event rating (levels 1–9, see table 1).
3. Identity of those involved (e.g. SN = sterile-nurse).
4. Time of event (in min).
5. Event description (a brief description of observed distraction or interruption).

To account for the variability in the occurrence of different types of event and to show how category types compared with one another, the data were analysed to show the following.

2.2.2. Data analysis.

2.2.2.1. Count and frequency.

1. The number of cases where particular events were recorded (n-case).
2. The mean count of events for each type from 'n-case' (mean-count).

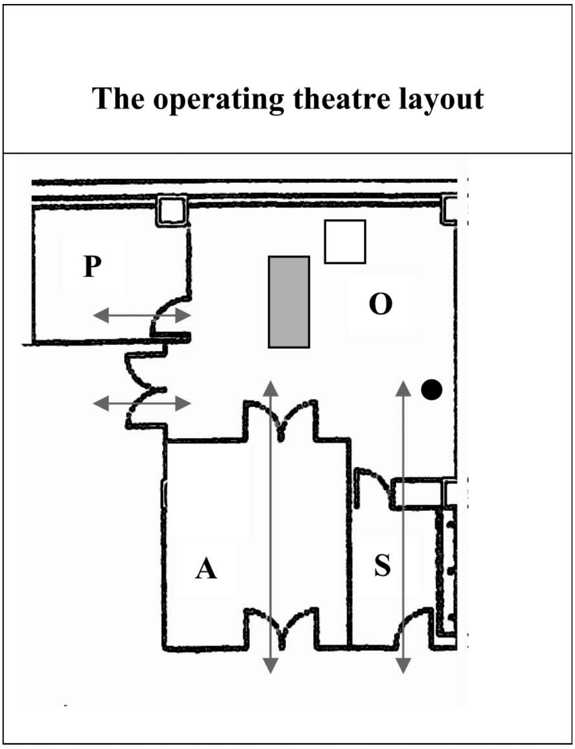

Figure 1. Geographic view of the operating theatre sampled, illustrating the flow of personnel in and out of the operating theatre. ↔ and ↕ indicate the bi-directional flow of movement through the theatre area. O = operating theatre; P = instrument preparation room; A = anaesthetic room; S = scrub room. ■ represents the operating table; □ represents the anaesthetic machine. This layout was observed in all 50 cases sampled. • indicates approximately where the observer was normally positioned during observation.

3. The maximum count of events from individual cases (max-count).
4. The total count of events from the 50-case sample (total-count).
5. The frequency of events per case proportional to case duration (mean-freq).

However, counts and frequencies may not show how much each type of event typically interferes with the team's work and how much it contributes to a sample of 50 cases. The data were therefore analysed to show the following.

2.2.2.2. Rating and interference.

1. The mean rating per event assigned to all events recorded (mean-rate).
2. The sum interference per case, summing all rated events per case (I-case).
3. The sum interference per event for the whole 50-case sample (I-sample).
4. The mean interference from each event from n-case (I-mean case).
5. The sum interference per case proportional to case duration (I-freq).

2.2.2.3. Statistical tests. Statistical tests included non-parametric Spearman's rho correlation testing the agreement between the two observers on their ratings of individual events and ratings summed for each case (I-case). Chi-square tests were also applied to the data from each category of interference, tabulated across disciplinary groups. Parametric tests included one-way ANOVA on interference data, comparing operation types and disciplinary groups. Pearson correlation analysis was applied to the data on interference, door-opening frequency and noise.

3. Results

3.1. *The operations*

The open operations sampled included anterior resection, sigmoid colostomy, appendectomy, central line insertion, formation of loop colostomy, gastrectomy, hemi-colectomy, hernia, small bowel resection, reversal of temporary ileostomy and perianal abscess drainage. The laparoscopic or minimal access operations included cholecystectomy, fundoplication, hernia and appendectomy. Intra-operative case duration ranged from 13 min to 217 min with a mean of 55.62 (SE $\pm$ 5.44).

3.2. *Inter-rater reliability*

A basic test of inter-rater reliability was the agreement between observers' overall summed rating of interference for each case that accounted for all observed case interference. The sum interference level was analysed for each of the 11 cases jointly measured. Spearman's rho correlation of ratings by ranked order showed a high level of correlation ($r_s(n=11)$ 0.89, $p < 0.001$); with an intra-class correlation coefficient of 0.85 ($F(n=11)=13.43$, $p < 0.001$). This confirmed a high positive agreement between observers' report of overall interference level. Of course, summed levels of interference for each case may hide considerable variation in the ratings between observers assigned to individual events. Therefore, from the 11 cases jointly measured, the first individual item from each category of interference was used to compare individual ratings. A total of 48 jointly rated events were identified from record sheets, cross-referenced by category and by the recorded time of the occurrence. Spearman's rho correlation analysis showed positive correlation between the ratings on specific events ($r_s(n=48)=0.60$, $p < 0.001$), with an intra-class correlation coefficient of 0.65 ($F(n=48)=4.746$, $p < 0.001$), reflecting reasonable agreement between observers' scores on individually rated items. For reporting the results, only the first set of the 11 jointly measured cases was used, as a second set was redundant after reliability testing.

3.3. *Sources of interference*

The total counts of events per case ranged from one to 39, with a mean of 13.56 (SE $\pm$ 1.12). The total count of events per case as a proportion of operative time ranged from 0.04 to 0.86 per min, with an average of 0.29 (SE $\pm$ 0.02). Figure 2 shows the total count of recorded events for each category from the entire 50-case sample. The integer count of observed events for each source category from the sample is shown by the dark spots, which correspond to the right axis of figure 2. The most frequent event type recorded was movement behind the video monitor during laparoscopic operations. The

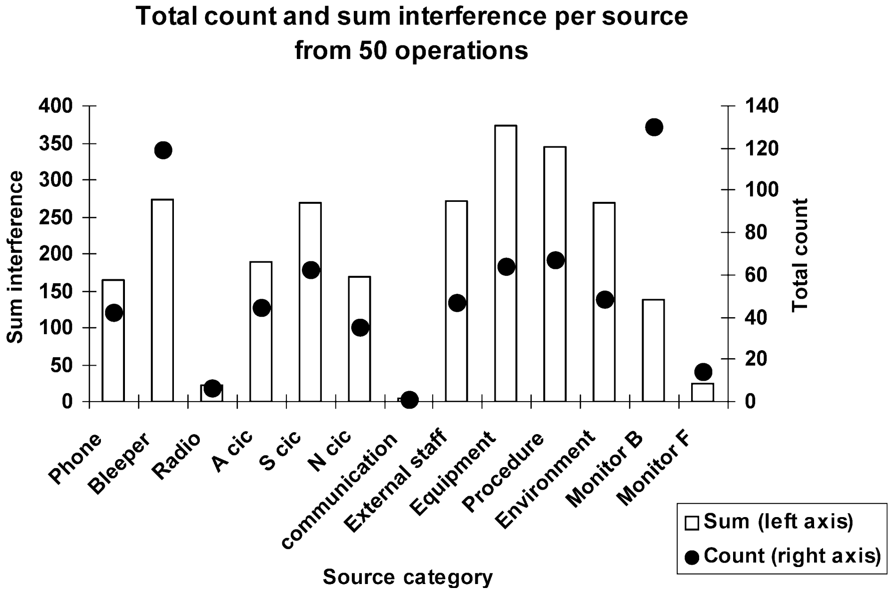

Figure 2. The integer count of observed events for each source category from the sample is shown by •, which correspond to the right axis. The interference rating of events for each source category is shown by the bars, which correspond to the left axis. A = anaesthetists; N = nurses; S = surgeons; Phone = any phone in theatre or next to theatre; Bleeper = any bleeper activated in theatre; Radio = action or response to the radio causing distraction. Case irrelevant communication (cic) = any conversation irrelevant to the case at hand. Comm = communication difficulties, e.g. lack of response to request. External staff = anyone not part of the team in theatre. Equipment = any item of equipment or provision not at hand or failing. Work environment = workspace and human–interface problems. Procedural = events intrinsic to the case not fitting other categories. Monitor-B = movement behind video display monitor. Monitor-F = movement in front of video display monitor.

least frequent sources were radio and communication exchange problems. The first column of table 2 shows the number of cases where each category of event was recorded (n-case). The mean and maximum counts of events from n-cases are shown in the second and the third columns, respectively.

Individual events may occur several times in a single case; repeated events are likely to have a cumulative effect. Therefore, the fifth column of table 2 (I-mean) shows the mean interference for each event type from the cases that were recorded (n-cases). The sum interference per case (I-case) from all categories, ranged from 6 to 124, with a mean of 50.14 (SE $\pm$ 3.94). The sum interference per case proportional to operative case duration (I-freq) ranged from 0.22 to 3.04 per min, with an average of 1.04 (SE $\pm$ 0.07). The last column of table 2 (I-sample) and the bars of figure 2 show the sum interference of each category for the whole 50-case sample, which correspond to the left axis.

Figure 2 therefore shows a pattern of results, where categories may be compared in their difference between count and sum interference, according to the scoring criteria. For instance, there were 119 bleeper events gaining an interference score of 274, whereas there were half as many external staff events (47) gaining a similar interference score of 271. Moreover, there were 64 equipment events scoring 374. The bars in figure 2 show that

Table 2. Summary data on recorded events from the 50-case sample, in the frequency and levels of interference for each source category.

Source	n cases*	Mean count†	Max count‡	Mean rate§	I – mean/case¶	I – sample‖
Phone	26	1.62	4	3.90	6.31	164
Bleeper	36	3.31	21	2.30	8.56	274
Radio	6	1.00	1	3.67	3.67	22
A cic	27	1.63	4	4.30	7.00	189
S cic	33	1.88	5	4.35	8.18	270
N cic	24	1.46	3	4.83	7.04	169
Communication	10	1.10	2	5.77	5.22	5
External staff	25	1.44	3	4.75	11.18	271
Equipment	35	1.83	7	5.84	10.69	374
Procedural	36	1.86	6	5.13	9.56	344
Environment	27	1.78	6	5.58	9.93	268
Monitor B	26	5.00	13	1.06	5.52	138
Monitor F	9	1.56	4	1.71	3.00	24

*Shows the number of cases where a particular source was recorded.
†Refers to the mean number of events from cases where they were recorded.
‡Shows the maximum count of recorded events for each source in a single case.
§Shows the mean rating assigned to events across the 50-case sample.
¶Shows the mean interference (I) from each category to cases where those events were recorded.
‖Shows the sum I of each category from the entire 50-case sample.
A = anaesthetists; N = nurses; S = surgeons; cic = case-irrelevant conversation; Monitor-B = movement behind video display monitor; Monitor-F in front of video display monitor.

case irrelevant conversation and external staff contributed considerably to overall interference. The third column (mean-rate) of table 2 also shows that the mean ratings for work environment, equipment and external staff were also the highest. Table 3 shows a breakdown of events recorded under the categories of equipment, work environment and procedure.

3.4. *Operation type*

Total counts proportional to operative case duration were also significantly higher for laparoscopic operations (0.38 per min, SE $\pm$ 0.03) compared to open (0.18 per min, SE $\pm$ 0.02) ($F(1,49) = 16.63$, $p < 0.001$). Interference proportional to operative case duration (I-freq) for laparoscopic operations (n = 29, mean = 1.18 per min, SE $\pm$ 0.08) was higher than open operations (n = 21, mean = 0.81 per min, SE $\pm$ 0.13). This difference reached significance in an independent one-way ANOVA on operating type (2) ($F(1,49) = 5.78$, $p < 0.05$). After subtracting the highly frequent monitor count, the difference of interference frequency between laparoscopic (0.25, SE $\pm$ 0.02) and open operations (0.18, SE $\pm$ 0.02) remained marginally significant ($F(1,49) = 3.72$, $p = 0.06$).

3.5. *Group relationship to interference*

Table 4 shows the total count of the different sources of interference for the three groups, namely, anaesthetic (A), nurse (N) and surgeon (S) and respective statistical values for categorical analysis. Overall, the surgeon group was more frequently distracted, with a total count of 276, compared to the nurse (213) and anaesthetist groups (116). Table 4

Table 3. Sample lists of events falling into the equipment, work environment and procedure categories.

Equipment	Work environment	Procedural
• Diathermy (electrosurgical coagulator) off or settings not correct • Instruments and provisions not at hand • Laparoscopic graspers not working • Sterilizer noise twice • Water hose detached • Failing table adjustment • Anaesthetic machine false-alarm • Faulty anaesthetic equipment – oxygen output • Suction not on • Endoscope not working • Clamp for abdomen jammed • Electric point not working	• Misting of camera lens • Spotlight needs redirecting • Poor image focus • Surgeon must over-extend to reach equipment • Image affected by diathermy use • Surgeon needs diathermy moved for stack space • Surgeon trips over floor cable • Surgeon must turn during laparoscopy to find diathermy pedals • Spotlight inadequate • Surgeon says the room is too warm • Sterile-nurse needs foot stool to view operative site	• Nurse unfamiliar with diathermy • Lead surgeon teaching students in theatre • Anaesthetist needs lights on to tend to patient • Surgeon leaves to attend to bleeper message • Anaesthetist leaves to getprovisions from anaesthetic room • No bloods ready for patient • Anaesthetist leaves to get provisions from anaesthetic room • Anaesthetist collects provisions from anaesthetic room

Table 4. A sum count of groups observed effected by separate sources of interference.

Source	A	N	S	χ^2*	$p <$
Phone	5	20	18	9.25	0.01
Bleeper	7	45	13	38.52	0.0001
Radio	2	4	3	–	NS
A cic	21	3	24	16.12	0.0001
S cic	4	10	58	58.62	0.0001
N cic	2	29	13	25.13	0.0001
Communication	1	5	7	–	NS
External staff	8	26	31	13.5	0.001
Equipment	12	40	47	20.78	0.0001
Procedural	42	10	18	23.77	0.0001
Work environments	12	21	44	21.22	0.0001
Group total	116	213	276		

*Test results show significant differences between groups, affected by different sources.
A = anaesthetists; N = nurses; S = surgeons; cic = case-irrelevant conversation.

also shows that with the exception of the radio and communication, groups were differentially effected by different events.

3.6. *Door opening and interference*

There were three doorways to the operating theatre (see figure 1); all were used as both entrance and exit to and from the operating theatre. Door opening counts per case ranged from 8 to 79, with an average of 33 (SD 16.26). The mean frequency of door opening as a proportion of operative case duration was 0.68 per min (SE $\pm$ 0.03). A one-way ANOVA showed that groups (anaesthetists, nurses and surgeons) differed significantly in their frequency of door-opening ($F(2,147) = 26.16$, $p < 0.001$). Post hoc tests (Tukey Least Significant Difference) confirmed that door-opening frequency was higher in the nurse group (0.33) compared to the surgeon group (0.11) (mean difference = 0.21, $p < 0.001$) and the anaesthetist group (0.25) (mean difference = 0.07, $p < 0.01$).

Many events, but not all, involved door opening. Door opening and interference level as a proportion of intra-operative case duration were both analysed with Pearson's correlation test. As illustrated in figure 3, the analysis showed a significant positive relationship between interference and door opening ($r(n = 50) = 0.47$, $p < 0.001$).

3.7. *Noise and interference*

The sound meter was not in use for five cases due to a software fault. The remaining sound pressure level data from each case was transferred to spreadsheet. The meter had been set to automatically log sound pressure levels every 2 s. That data were used to calculate summary statistics for each case. The absolute minimum level of noise recorded was 39.20 dB(A) and the absolute maximum was 94.40 dB(A). The minimum mean noise level per case was 54.40 dB(A) and the maximum was 62.94 dB(A), with an overall mean of 57.80 (SE $\pm$ 0.29). Pearson correlation analysis did not show a significant positive relationship between interference and noise ($r(n = 45) = 0.25$, $p = 0.09$), but a larger sample might have gained a significant result.

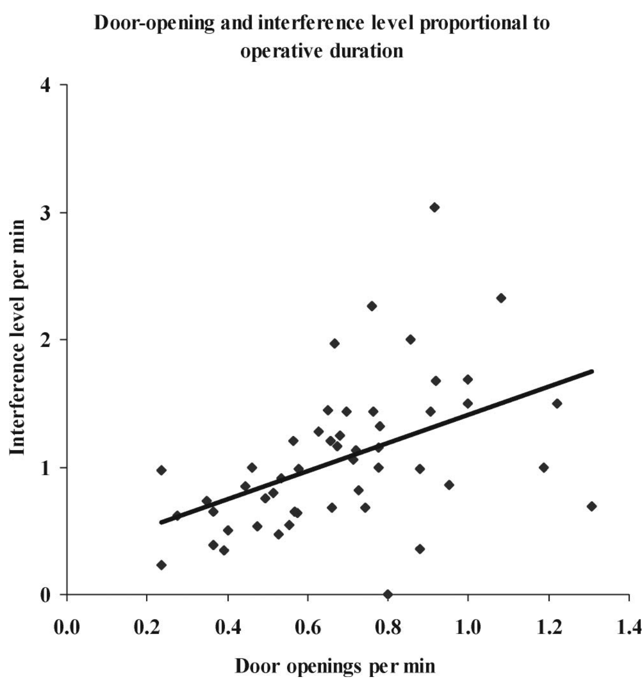

Figure 3. Theatre door opening proportional to operative duration plotted against the sum interference level proportional to operative duration. The significant positive correlation suggests that door opening is indicative of interference level.

4. Discussion

4.1. *The main findings*

A sample of 50 surgical cases from a single operating theatre showed a high frequency of distraction and interruption. The total count of events recorded per case ranged from one to 39, with a mean rate of 0.29 per min. The highest frequency events included bleepers in the operating theatre 21 times in one operation; movement behind the video monitors that guided surgical action; and conversations among personnel that were unrelated to the case at hand. Procedure, equipment and environment events were less frequent than the aforementioned, but they often involved several team members.

To account for the level of interference that events might have on the work of surgical teams, observed events were also rated using the scale in table 1. To obtain a measure of interference, all ratings from each case were summed. The sum interference per case that accounted for all events rated ranged from six to 124. Longer operations would naturally experience more events than shorter operations, all things being equal. Therefore, calculating the sum interference as a proportion of operative duration is important to identify cases where interference might be particularly high or intense. Interference ranged from 0.22 to 3.04 per min, with a mean of 1.04 per min. A high level of interference derived from sources irrelevant to the case, from bleepers, telephones, conversations and external staff. While bleepers and movement around video monitors were highly frequent, they appeared less interfering because few team members appeared distracted by them. On the other hand, circulating nurses often dealt with bleepers, personal telephone calls and theatre telephone calls in parallel to supporting the sterile-nurse and surgeons. Therefore, extraneous sources of interference were likely to have a disproportionate observed effect on the nursing group compared to others.

Together, equipment, work environment and procedural events accounted for a considerable amount of interference. Equipment problems, work environment difficulties and provisions not being at hand in the operating theatre often involved all groups, but mostly the surgeons and sterile-nurses. In some cases, equipment problems interrupted surgical teamwork because equipment had not been fully prepared and its functioning had not been checked pre-operatively. Equipment and work environment problems were more frequent in laparoscopic operations, contributing to a significantly higher interference for laparoscopic operations compared to open operations. This is perhaps unsurprising as laparoscopic procedures involve more complex technology than open procedures.

Figure 1 illustrates the layout of the operating theatre and adjacent rooms where equipment and provisions for surgery and anaesthetics were usually stored. The location of equipment and provisions and the journey lengths to collect them exacerbated intra-operative interference. Bleepers and phone calls also demanded that circulating nurses leave the operating theatre and external staff often entered and exited the operating theatre. In some cases, the operating theatre space appeared to function as a thoroughfare. Consequently, operating theatre doors opened twice every 3 min on average. Intuitively, door-opening frequency was indicative of interference level in theatre, evidenced by the positive correlation found between the two measures.

The observed flow of personnel through theatre suggested a general lack of control over different zones of activity. This argument is supported by the fact that it was noted that all eight theatres in the hospital, including the one sampled, had 'theatre-in-use' lights on their external walls, none of which was ever seen to be used. The theatre lights were positioned above the entrance to the scrub-room that is illustrated in figure 1. It is noteworthy that modern operating theatres are installed with complex infrastructure and technology to ensure that operating theatres maintain a system of ultra-clean air. These air conditioning systems function to reduce the likelihood of airborne agents that may cause infection; a function compromised by excessive door opening.

The finding of considerable case-irrelevant conversation and deviation from noise criteria reflects the events described. The overall case mean for noise was 57.80 (SE $\pm$ .29), reaching 62.94 dB(A) for one case. Basic ambient noise level is an important factor in task performance (e.g. Gillie and Broadbent 1989). For satisfactory speech intelligibility, there should be a 10dB(A) difference between the ambient noise and speech noise at the position of the speaker (Chartered Institute of Building Services Engineers 1994). Therefore, if relaxed-to-raised vocal effort is up to 66dB(A), the ambient noise level in theatre should be no more than 56dB(A) for communication within a 1–2 m range. Therefore, noise levels in the operating theatre sampled were acceptable in some but not in all cases. This confirms previous research findings that noise is a potential problem for performance in operating theatres (Shapiro and Berland 1972), as it is in other work environments where concentration is paramount to task performance (Banbury and Berry 2005).

There was a match in the frequency of recorded events (0.29 per min) between this study and that of Chisholm *et al.*'s (2000) study, despite the difference in contexts, namely, surgery and accident-emergency, and in the events recorded. This similarity in results may be due to chance or it may reflect a similarity of events across health-care settings. Alternatively, the similarity in results might have derived from an aspect of the observation process itself, perhaps reflecting the limitation of what an observer might feasibly observe in a given situation and in a certain period of time.

4.2. *The method*

Recording multiple concurrent events was particularly difficult for the observers, but deemed manageable. One particular problem was in categorizing events as they occurred. Some events could be coded in more than one category, but using the initial stimulus of a distracting event as the criteria for categorization seemed to resolve that problem satisfactorily. As an alternative, the observations might have been recorded and then rated and categorized retrospectively, post-observation. On the other hand, retrospective coding might feasibly increase bias in the categorization process, where inference rather than observation might bias categorization. Given the main problem of the load on observation in this study, it seems a logical step to perhaps confine observation to fewer categories. The ordinal scale could also be reduced to fewer points or more strict criteria applied to the allocation of scale points to events. Inter-rater agreement on the levels of interference was acceptable, particularly in discriminating overall interference per case. However, the agreement on individual events could be improved with further refinement of the measures.

Observation in the operating theatre was accepted by theatre personnel in nearly every case. Future observational research might instead utilize operating theatre cameras to measure the effects of distracting events remotely and separately from observation *in situ*. Video recorded cases would be useful for detailed analysis of events. Subjective ratings of events in the operating theatre by those directly involved could be related to observational measures. Controlled experiments (e.g. Einstein *et al.* 2003) might also be useful in gaining further understanding of the effects of specific distraction and interruption on individual performance.

However, when the effects of distraction and interruption are diffuse, cumulative and interactive, neither experiments nor individual members may detect or describe them effectively. The position and duration of distracting and interrupting events and their combinations and interactions are difficult to view as a participant in the work activity and reductive experimentation is clearly inappropriate. Attempts to measure distraction and interruption in relation to their effect on the team *in situ* are, therefore, important for researching the environment that teams create.

The authors chose to form a measure of interference based upon the relative involvement of the team because counts and frequencies do not show anything about the level of interference experienced in the operating theatre. However, it was noted that in some operations many events overlapped and produced an intensity of interference that this study does not fully describe. It will be important to attempt to measure intensities and variation of interference in future work and to make more use of the temporal or sequential aspects of the recorded events.

The method reported here also assumes that interference is greater the more team members are involved in an event. While that is a logical basis for measurement, it discounts the fact that some individual tasks may be more sensitive to interference than other tasks. For instance, when a surgeon is distracted during incision there is likely more probability of surgical outcome being affected than when a circulating nurse is distracted from unwrapping provisions. The measure was focused, indeed biased, toward surgical action, but did not account for such qualitative difference.

There are inevitable trade-offs in observational sampling, between focus and scope and between other factors, such as usability, data quality and quantity. From a systems perspective it is important to consider the performance of all team members and address surgical teams as singular units of analysis, rather than focus upon individuals alone. In order to address the interactions in surgical teams it is necessary to view the system with sufficiently broad scope.

For instance, it was noted that a rather informal structure existed for dealing with interference in the operating theatre sampled. Circulating nurses providing backup to the team tended to deal with bleepers, theatre telephone calls and other external staff, acting as a barrier or filter of information flowing into theatre. However, in dealing with such interference the primary work of the nurses in supporting sterile-nurses and surgeons was evidently compromised. From the surgeon's view, impaired intra-operative team performance might be attributed to nursing competence alone, rather than the conditions of their work and workload, which may serve to perpetuate conflict and stress in surgical teams. Indeed, evidence has shown that there is a culture in the health-care domain of denying causes of stress and fatigue and a general lack of openness to discuss systemic problems (Sexton *et al.* 2000).

It is necessary to consider the work environment and work distribution in surgical teams more thoroughly. Not all breaks from task activity are detrimental to performance; indeed, they may be necessary to enhance performance. Some conversation may reduce social tension or boredom and may help maintain heightened awareness and vigilance (Weinger and Englund 1990, Quintus and George 2003). Addressing the issues concerning team performance in surgery is not an easy task for the groups who work in the operating theatre. To describe and measure the pattern of interference that surgical teams experience is a useful step forward in addressing those concerns.

5. Conclusions

This study shows that observational measurement of distraction and interruption in the operating theatre is feasible. The measures need further validation by obtaining agreement to other relevant measures. However, on balance, the method used to measure and describe interference in the operating theatre was effective and reasonably reliable. The initial results suggest that from one sample of the operating theatre environment there is considerable distraction and interruption in the operating theatre interfering with the work of surgical teams that may impact on team performance and surgical outcome. Factors intrinsic to operative procedures, in failing or missing equipment or in difficulties with the work environment, distracted and interrupted surgical teams from their work. Factors extrinsic to the case at hand, including bleepers, phone calls and external staff entering the operating theatre also interfered with their work. There was some convergence on the different measures obtained, in that door opening and interference correlated. In short, the results obtained are a useful description of the operating theatre environment. The results suggest that the operating theatre environment should be the subject of further study.

Surgery is a high-risk activity. Therefore, it would naturally be expected that there would be a high level of control in the operating theatre environment, a high level of reliability in equipment and a low level of interference overall. It remains to be seen how much interference is experienced in different operating theatres and what effect it has on the performance of surgical teams and patient care.

Acknowledgements

BUPA Foundation, BUPA House, 15–19 Bloomsbury Way, London WC1A 2BA, http://www.bupafoundation.com

Smith and Nephew Foundation, 15 Adam Street, London WC2N 6LA, http://www.snfoundation.org.uk/

Patient Safety Research Program, The Department of Health, Richmond House, 79 Whitehall, London SW1A 2NL, http://www.dh.gsi.gov.uk

References

Banbury, S. and Berry, D.C., 2005, Office noise and employee concentration: Identifying causes of disruption and potential improvements. *Ergonomics*, **48,** 25–37.

Brennan, T.A., Leape, L.L., Laird, N.M., Hebert, L., Localio, A.R., Lawthers, A.G., Newhouse, J.P., Weiler, P.C. and Hiatt, H.H., 1991, Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *New England Journal of Medicine*, **324,** 370–376.

Chartered Institute of Building Services Engineers, 1994, *Environmental Criteria for Design: Guide A* (London: CIBSE).

Chisholm, C.D., Collison, E.K., Nelson, D.R. and Cordell, W.H., 2000, Emergency department workplace interruptions: are emergency physicians 'interrupt-driven' and 'multitasking'? *Academic Emergency Medicine*, **7,** 1239–1243.

Chou, C.D. and Funk, K., 1990, Management of multiple tasks: Cockpit task management errors. In *IEEE International Conference on Systems, Man, and Cybernetics*, November 1990, Los Angles, CA. pp. 1154–1156.

Department of Health, 2000, An organisation with a memory report of an expert group on learning from adverse events in the NHS. Available online at: http://www.dh.gov.uk/assetRoot/04/06/50/86/04065086.pdf

Dismukes, R.K., Loukopoulos, L.D. and Jobe, K.K., 2001, The challenges of managing concurrent and deferred tasks. In *Proceedings of the 11th International Symposium on Aviation Psychology*, R. Jensen (Ed.) (Columbus, OH: Ohio State University).

Einstein, G.O., McDaniel, M.A., Williford, C.L., Pagan, J.L. and Dismukes, R.K., 2003, Forgetting of intentions in demanding situations is rapid. *Journal of Experimental Psychology: Applied*, **9,** 147–162.

Gerbrands, A., Albayrak, A. and Kazemier, G., 2004, Ergonomic evaluation of the work of the Scrub Nurse. *Minimally Invasive Therapy and Allied Technologies*, **13,** 142–146.

Gillie, T. and Broadbent, D.E., 1989, What makes interruptions disruptive? A study of length, similarity and complexity. *Psychological Research*, **50,** 243–250.

Healey, A.N., Undre, S. and Vincent, C.A., 2004, Developing observational measures of teamwork in surgery. *Quality and Safety in Healthcare*, **13,** Suppl. 1, 33–40.

Latorella, K.A., 1996, Investigating interruptions: Implications for flight deck performance. Doctoral dissertation, State University of New York, Buffalo. (Also published in 1999 as NASA Technical Memorandum 209707, National Aviation and Space Administration, Washington, DC.)

Loukopoulos, L.D., Dismukes, R.K. and Barshi, I., 2001, Cockpit interruptions and interferences: A line observation study. In *Proceedings of the 11th International Symposium on Aviation Psychology*, R. Jensen (Ed.) (Columbus, OH: Ohio State University).

Quintus, J.R. and George, J.M., 2003, Work interrupted. A closer-look at the role of interruptions in organisational life. *Academy of Management*, **28,** 494–508.

Reason, J., 2000, Human error: models and management. *British Medical Journal*, **320,** 768–770.

Sexton, J.B., Thomas, E.J. and Helmreich, R.L., 2000, Error, stress and teamwork in medicine and aviation: cross sectional surveys. *British Medical Journal*, **320,** 745–749.

Shapiro, R.A. and Berland, T., 1972, Noise in the operating room. *New England Journal of Medicine*, **287,** 1236–1238.

Van Veelen, M.A., Nederlof, E.A.L., Goosens, R.H.M., Schot, C.J. and Jakimowicz, J.J., 2003, Ergonomic problems encountered by the medical team related to products used for minimally invasive surgery. *Surgical Endoscopy*, **17,** 1077–1081.

Weinger, M.B. and Englund, C.E., 1990, Ergonomic and human factors affecting anesthetic vigilance and monitoring performance in the operating room environment. *Anesthesiology*, **73,** 995–1021.

Welty, G., Schippers, E., Grablowitz, V., Lawong, A.G., Tittel, A. and Schumpelick, V., 2002, Is laparoscopic cholecystectomy a mature technique? *Surgical Endoscopy*, **16,** 820–827.

Wilson, J.R., 2000, Fundamentals of ergonomics in theory and practice. *Applied Ergonomics*, **31,** 557–567.

Zijlstra, F.R.H., Roe, R.A., Leonova, A.B. and Krediet, I., 1999, Temporal factors in mental work: effects of interrupted activities. *Journal of Occupational and Organisational Psychology*, **72,** 163–186.