BMJ Open

Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016977
Article Type:	Research
Date Submitted by the Author:	24-Mar-2017
Complete List of Authors:	Tarling, Maggie; City University School of Health Sciences, Nursing Jones, Anne; Kings College London, Adult Nursing Murrells, Trevor; Kings College London, National Nursing Research Unit McCutcheon, Helen; University of Queensland, Faculty of Health and Behavioural Sciences
Primary Subject Heading:	Nursing
Secondary Subject Heading:	Communication, Health services research, Intensive care, Qualitative research
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Adult intensive & critical care < INTENSIVE & CRITICAL CARE, QUALITATIVE RESEARCH

Title: Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Author / Corresponding author:

Maggie Tarling
School of Health Sciences
City, University of London
Northampton Square
London, EC1 0HB
United Kingdom

Email: Maggie.Tarling.2@city.ac.uk
T: +44 (0)20 7040 0201

Co-Authors:

Anne Jones
Florence Nightingale Faculty of Nursing & Midwifery
King's College London
57 Waterloo Road
London, SE1 8WA
United Kingdom

Trevor Murrells
Florence Nightingale Faculty of Nursing & Midwifery
King's College London
57 Waterloo Road
London, SE1 8WA
United Kingdom

Helen McCutcheon
The University of Queensland
Brisbane St Lucia, QLD 4072
Australia

Keywords: patient safety; safety climate; nursing; mixed methods; human factors

Word count: 4680

ABSTRACT

Objectives: To explore the underlying factors that influences the presence of safety climate in nursing practice.

Design: A sequential mixed methods design included a cross-sectional survey using the Safety Climate Questionnaire (SCQ) and thematic analysis of focus group discussions. Confirmatory Factor Analysis (CFA) was used to validate the factor structure of the SCQ. Factor scores were compared between nurses working in operating theatres, critical care and ward areas. Results from the survey and the thematic analysis were then compared and synthesised.

Setting: A London University.

Participants: 319 registered nurses working in acute hospital settings completed the SCQ and a further 23 nurses participated in focus groups.

Results: CFA indicated that there was a good model fit on some criteria (χ^2 = 1683.699, df 824, $p < 0.001$; χ^2/ df = 2.04; RMSEA = 0.058) but a less acceptable fit on Comparative Fit Index (CFI) = 0.804. There was a statistically significant difference between clinical specialisms in Management Commitment (F [4,266] = 4.66, p = 0.001). Nurses working in operating theatres had lower scores compared with ward areas and they also reported negative perceptions about management in their focus group. There was significant variation in scores for Communication across clinical specialism (F [4,266] = 2.62, p = 0.035) but none of the pair-wise comparisons achieved statistical significance. Thematic analysis identified themes of Human Factors, Clinical Management and Protecting Patients. The System and the Human Side of Caring was identified as a meta-theme. There were areas of overlap and differences between the dimensions of the SCQ and the findings of the thematic analysis.

Conclusions: The results suggest that the SCQ has some utility but requires further exploration. The findings indicate that safety in nursing practice is a complex interaction between safety systems and the social and interpersonal aspects of clinical practice.

ARTICLE SUMMARY

Strengths and limitations of this study:

- The results of the study indicate that there is an important and complex link between human factor approaches used in nursing practice and the interpersonal aspects of care.
- This work makes a unique contribution to understanding safety climate in nursing practice in the UK setting.
- The Confirmatory Factor Analysis of the Safety Climate Questionnaire indicated that the model fit could be improved but further psychometric exploratory analysis may be warranted.
- The results need to be considered in the light of a cross-sectional survey response rate of 57% and a low number of participants in some of the focus groups.

INTRODUCTION

There is a growing consensus in healthcare safety research that organisational culture is critical for patient safety [1] and that safety management should move away from depending on lagging indicators of safety issues, such as incident reports, and move towards leading indicators, such as, measures of safety climate.[2] Patient safety culture is defined as aspects of organisational culture that are 'the product of individual and group values, attitudes, perceptions, competencies and patterns of behaviour that determine the commitment to and the style and proficiency of an organisations' health and safety management'.[3] Safety climate is defined as a measurable feature of staff's attitudes and perceptions of an organisations underlying safety culture at any point in time.[4] There is evidence that safety climate is open to change and has an impact on individual safety behaviour and an important factor in improving patient safety.[5,6]

The Safety Climate Questionnaire (SCQ) developed in the UK [7] has been used extensively in the NHS by the Royal College of Nursing.[8] However the SCQ was originally developed for use in the UK petroleum industry as part of a tool kit to measure safety climate. The SCQ measures nine factors that contribute to safety climate, namely Management commitment, Communication, Priority of safety, Safety rules and Procedures, Supportive environment, Involvement, Personal priorities and need for safety, Personal appreciation of risk and Work environment.[9] It is noted that the petroleum industry exhibits aspects of a High Reliability Organisation,[10] defined as 'organisations that are able to manage and sustain almost error-free performance despite operating in hazardous conditions where the consequences of errors could be catastrophic'[11] and as such lessons learnt from High Reliability Organisations have underpinned developments in safety and risk management in the NHS.[12] The petroleum industry is a very different setting from healthcare organisations but it is possible that their safety management systems could provide beneficial outcomes in safety and risk management in the healthcare setting.[13] Pilot testing of the SCQ undertaken within the NHS focused upon how acceptable and useable the tool was.[14] However, a confirmatory factor analysis of the tool was not undertaken to validate its psychometric properties with a healthcare population.

Research evidence suggests that measures of safety climate vary between and within healthcare organisations and that there is limited understanding of the factors that may influence and explain the sources of these variations.[15] Several research studies have reported safety climate scores varying across different clinical specialities with some reporting less safe climates in operating theatres, critical care and emergency departments compared to surgical and medical inpatient areas [16-18] and others reporting a safer climate in critical care [19-20]. However, none of this research has been undertaken in the UK. The underlying reasons for these variations in safety climate are unclear at the present time. Understanding the underlying factors that influence healthcare practitioner's perceptions of safety climate is important for the development of strategies to improve patient safety.

As a subset of healthcare practitioners, nurses make an important contribution to patient care and evidence indicates that nurse-staffing levels have a direct impact on patient mortality. [21-22] Therefore it is important to understand how nurses perceive safety climate as this may have a direct impact on patient safety. This mixed methods study set out to explore the underlying factors that contribute to safety climate in nursing practice. The aims of the study were to investigate the factor structure of the SCQ, determine whether there are differences in the perception of safety climate between nurses working in critical care, operating theatres, surgical and medical wards in acute hospital settings in the UK and understand the meaning that nurses working in these different clinical settings attribute to their understanding of patient safety.

METHOD

Design

The study design was a fully mixed, sequential, equal status, mixed methods design and was conducted in two phases.[23] The first phase of the study measured and then compared safety climate scores between groups of nurses working in operating theatres, critical care, surgical and medical ward areas. As the factor structure of the SCQ had not been evaluated in a nursing sample a Confirmatory Factor Analysis (CFA) was also undertaken. The results from the cross-sectional survey were used to structure the focus group discussions held with groups of nurses from operating theatres, critical care and ward areas. The results of both phases of the study were then jointly summarised in a statistics-by-theme format to facilitate

more in-depth inferences in order to consider potential mechanisms underlying safety climate.[24-25]

Setting and sample

Following local ethical approval participants were recruited from a qualified nursing population who attended a university that recruited from a wide range of NHS Trusts and private hospitals in the region. In the UK Band 5 and 6 nurses are qualified nurses who deliver bedside care. They were specifically chosen, because they have a direct impact upon patient care and safety in their everyday practice. A convenience sampling method was used and participants were approached by the researcher at the beginning of a teaching session and the purpose of the survey was explained. Information sheets were included with the questionnaire and completion of the questionnaire implied consent. All questionnaires distributed were collected at the end of the afternoon teaching session. The aim was to collect at least 300 questionnaires as this is considered by some to be the minimum number required for robust factor analysis.[26]

Measures and variables

A paper version of the SCQ was distributed to participants. Additional questions were added to the questionnaire in order to facilitate a stratified analysis to compare scores between nurses working in different clinical settings and measure potential factors that may influence perceptions of safety. These additional questions collected data on the clinical area the participant worked in, including whether they worked in a surgical ward, medical ward, critical care unit, operating theatre or other acute hospital unit. Further information included how long they had worked in their present position, how long they had worked in the speciality, how long they had been qualified and whether they had safety training and further training in their speciality. Participants were also asked to describe the type of training they had undertaken.

Data collection and management

The SCQ has 43 questions with a 5-point Likert scale response and is scored by allocating a value of 5 to the 'strongly agree' response, 4 to 'agree' response, 3 to the 'neither agree nor disagree' response, 2 to the 'disagree' response and 1 to the 'strongly disagree' response. The negative worded questions were allocated a reverse score by subtracting the initial score from 6. The initial scores from the questionnaires provided raw scores and these were transferred

into an Excel® 2013 spreadsheet. In order to ensure that the data entry was as accurate as possible a double data entry procedure was followed as recommended by Elliot et al.[27]. The Excel spreadsheet was then transferred into SPSS® V21 and a Little's 'missing completely at random' (MCAR) test was undertaken to ensure that any missing data was not introducing bias into the analysis.[28]

Data analysis

A confirmatory factor analysis of the SCQ scores was undertaken using SPSS® Amos V21. The original nine factor structure as identified by Cox and Cheyne was used as the *a priori* model to be confirmed by the factor analysis.[7] The following goodness of fit indices were used to test the model. Chi Square (χ^2) and the χ^2/df ratio, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The χ^2/df ratio overcomes the problem of a statistically significant χ^2 result associated with a larger sample sizes. A value of between 2 -3 is deemed as being acceptable the smaller value the better the fit.[29] The Comparative Fit Index (CFI) measures the difference in the non-centrality estimates of the baseline and proposed model with values ranging from 0 to 1. A cutoff value above 0.9 is considered to be an indication of a good model fit. The Root Mean Square Error of Approximation (RMSEA) measures the discrepancy between the hypothesised model and the population covariance matrices, and values range from 0 – 1. A RMSEA of less than 0.06 is indicative of an acceptable model fit with a recommended upper limit of 0.07.[30-31]

Once the CFA had been undertaken comparisons of safety climate dimensions (factors) were made between different clinical settings. Higher mean scores indicate a good safety climate. Dimension scores were compared between clinical specialisms using a general linear model (GLM) that adjusted for the following characteristics: years in current position, years qualified, years in specialism, specialist qualification and safety training. Adjusted means with 95%confidence intervals were calculated. Where there were differences between clinical specialism, based on the GLM F statistic, Bonferroni post-hoc pair wise comparisons were performed.

A Levene test of homogeneity of variance was conducted and residual plots produced, to ascertain whether the assumptions underpinning GLM had been met. A wild Bootstrap analysis was undertaken on the 'Personal priorities and need for safety' dimension to assess

whether non-equality of variance had biased the results. [32] The results remained very similar and only those from the GLM have been reported.

Following the survey a total of 23 nurses were recruited and participated in four focus groups (Operating theatre group = 8, Critical Care group = 9, ward A group = 3 and Ward B group = 3). All participants gave consent to participate in the focus groups. These participants had not participated in the survey and therefore had not completed the SCQ. The main topics covered were, their overall understanding of safety climate or culture, communication and manager commitment to safety, as these dimensions of safety climate had been found to be different between groups in the first phase of the study. Each focus group was facilitated by one researcher who acted as facilitator, and an observer who noted group dynamics and timed the session. The groups lasted between 40 to 50 minutes and were recorded. A six phase approach to a thematic analysis was undertaken [33]

RESULTS

Survey results

A total of 563 questionnaires were distributed and 319 questionnaires were completed and returned (response rate = 57%). Four questionnaires were excluded from the final analysis because they were completed by nurses who did not fulfil the selection criteria, i.e., not a band 5 or 6 adult nurse working in an acute hospital setting. Little's MCAR indicated that the missing data were missing completely at random and were unlikely therefore to unduly affect the results (Little's MCAR test: $\chi^2 = 2368.11$, df = 2292, p = 0.131)

Table one illustrates the demographic data of the participants according to the specialist areas they worked in. There were more participants from critical care units than from other groups. The group identified as other included participants who stated that they worked in acute hospital setting areas such as, out patients, care of the elderly, oncology and haematology. The numbers of participants in these areas was low so these were grouped together.

Table One: Mean and standard deviation (SD) χ^2 and for demographic data for critical care, operating theatres, medicine, surgery and other clinical areas.

	Critical Care (n = 107)	Operating theatres (n = 49)	Medicine (n = 70)	Surgery (n= 54)	Other (n=24)	χ^2
Present position Mean (SD) years	3.12 (2.60)	4.26 (3.58)	3.69 (3.35)	3.12 (2.22)	3.21 (2.48)	p = 0.442
Years qualified Mean (SD) years	7.63 (5.47)	8.90 (6.85)	8.14 (6.46)	6.93 (6.00)	8.60 (6.04)	p = 0.317
Years Specialism Mean (SD) years	4.30 (3.77)	6.34 (5.25)	5.02 (3.84)	4.29 (3.74)	4.13 (2.70)	p = 0.195
Specialist qualification Percentage	50% (54/107)	43% (20/49)	37% (26/70)	33% (18/54)	58% (14/24)	P = 0.029*
Safety training Percentage	71% (76/107)	55% (27/49)	69% (48/70)	59% (32/54)	67% (16/24)	P = 0.032*

* statistically significant difference

Across the groups the participants had been working in their present position between 3 to 4 years. There was more variability across the groups in terms of how long the participants had been qualified with the critical care and surgery ward nurses being qualified as a registered nurse for less time. There was some variation in the amount of time the participants had been working within the specialism and the results indicate that the participants had been working in other areas before finally working within their specialist areas. The percentage of those reporting having undergone safety training ($\chi^2 = 6.12$, df = 4, p = 0.032) and those participants reporting having a specialist qualification ($\chi^2 = 9.83$, df = 4, p = 0.029) varied significantly across clinical specialism. All other variables did not vary significantly across clinical specialism. All participants who had reported undergoing safety training undertaken in UK hospitals on an annual basis described this as mandatory training. Typically this includes training in manual handling, resuscitation and infection control.

Confirmatory Factor Analysis

The CFA goodness of fit measures indicated that there was a good model fit on some criteria with a significant Chi Square test ($\chi^2 = 1687.560$, df = 824, p = < 0.001). Both the χ^2/df ratio of 2.05 and RMSEA value of 0.058 (90% CI interval 0.054 to 0.062) indicated a good model

fit. However, the CFI was 0.805, although this was towards the higher end of the CFI range (0 to 1) it was below the acceptable threshold level (CFI >0.9) and suggests that the model could be improved.

The CFA regression weights (factor loadings) were similar to those from the original petroleum industry study. However, there were four items that were particularly low and related to the dimensions of supportive environment, personal appreciation of risk and work environment. In relation to a supportive environment the item relating to, 'A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate', had a regression weight of 0.150 and the item relating to, 'When people ignore safety procedures here I feel it is none of my business', had a regression weight of 0.291. In the dimension of personal appreciation of risk, the item, 'I am rarely worried about being injured in the job', had a regression weight of 0.110 and in the dimension of Work environment the item, 'This is a safer place to work than other Trusts I have worked for', had a regression weight of 0.270. These items may not make a significant contribution to the perception of safety climate in a nursing population. Cox and Cheyne [7] kept lower regression weighted items in their original questionnaire and suggested that these items should be used with caution.

Comparison of safety climate scores

Following the CFA the factor scores derived from the survey were used to go onto explore differences in safety climate scores between nurses working in different clinical specialisms. Comparisons were made between nurses working in critical care areas, operating theatres, medical wards, surgical wards and other acute hospital settings as described above. Table two shows the adjusted GLM mean, 95% confidence interval by clinical specialism and F statistic, for all of the safety climate dimensions. Overall the scores were towards the higher range on the safety climate scale and suggested that participants reported a fairly positive safety climate for most of the dimensions. However, the work environment factor had lower scores across all the groups whilst personal priority of safety scored highly across all groups. There was a statistically significant difference between groups for Management Commitment ($F [4,266] = 4.66$, $p = 0.001$) and for Communication ($F [4,266] = 2.62$, $p = 0.035$).

Table Two: Comparison of the nine safety climate dimensions across clinical specialism adjusting for profile variable.

m = significantly different from Medicine; s= significantly different from surgery; o= significantly different from operating theatres

		Critical Care	Operating theatres	Medical wards	Surgical wards	Other	F test, p
Management Commitment	Mean	3.48	3.27[m,s]	3.75[o]	3.66[o]	3.31	F (4,266) = 4.66, p= 0.001
	(95% CI)	(3.34, 3.62)	(3.07, 3.67)	(3.59, 3.91)	(3.47, 3.85)	(2.99, 3.63)	
Priority of safety	Mean	3.54	3.44	3.73	3.50	3.61	F (4,266) = 1.29, p = 0.27
	(95% CI)	(3.39, 3.69)	(3.22, 3.66)	(3.55, 3.91)	(3.30, 3.71)	(3.26, 3.96)	
Communication	Mean	3.19	3.17	3.50	3.35	3.13	F (4,266) = 2.62, p= 0.035
	(95% CI)	(3.04, 3.33)	(2.96, 3.38)	(3.33, 3.67)	(3.15, 3.54)	(2.79, 3.47)	
Safety rules	Mean	3.18	3.23	3.43	3.40	2.90	F (4.266) = 1.96, p = 0.10
	(95% CI)	(3.01, 3.36)	(2.98, 3.48)	(3.22, 3.64)	(3.17, 3.64)	(2.49, 3.31)	
Supportive environment	Mean	3.66	3.67	3.86	3.75	3.63	F (4,266) = 1.85, p = 0.12
	(95% CI)	(3.55, 3.76)	(3.51, 3.82)	(3.73, 3.98)	(3.60, 3.89)	(3.38, 3.88)	
Involvement in safety	Mean	3.31	3.45	3.50	3.63	3.37	F (4,266) = 1.87, p = 0.12
	(95% CI)	(3.16, 3.46)	(3.24, 3.66)	(3.33, 3.68)	(3.43, 3.82)	(3.03, 3.71)	
Personal priorities and need for safety	Mean	4.20	4.31	4.37	4.33	4.11	F(4,266) = 1.89, p = 0.11
	(95% CI)	(4.10, 4.30)	(4.16, 4.45)	(4.25, 4.48)	(4.20, 447)	(3.88, 4.34)	
Personal appreciation of risk	Mean	3.19	3.15	3.36	3.44	3.35	F (4,226) = 0.92, p = .080
	(95% CI)	(3.05, 3.32)	(2.96, 3.34)	(3.20, 3.52)	(3.26, 3.61)	(3.04, 3.65)	
Work environment	Mean	2.62	2.65	2.68	2.82	2.85	F (4,266) = .092, p = 0.45
	(95% CI)	(2.47, 2.77)	(2.44, 2.86)	(2.50, 2.85)	(2.62, 3.02)	(2.51, 3.20)	

A Bonferroni post-hoc test revealed that there was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses (Mean = 3.27, 95% CI 3.07 – 3.47), compared with nurses working in medical wards (Mean = 3.75, 95% CI 3.59 – 3.91) and surgical ward settings (Mean = 3.66, 95% CI 3.47 – 3.85). Although there was significant variation in safety climate scores for communication across clinical specialism, none of the pair-wise comparisons achieved statistical significance at the 5% level, although the difference between critical care (Mean = 3.19, 95% CI 3.04 – 3.33) and the medical wards (Mean = 3.50 95% CI 3.33 – 3.67) came close (p = 0.056).

Thematic analysis

The results of the cross-sectional survey indicated a difference between nurses on the dimensions of Management commitment and though not statistically significant, Communication. During the focus groups participants were invited to discuss their understanding of safety culture and for their views of management and communication related to safety. Specific details of the differences found in the survey were not disclosed to the participants in order not to lead the discussion. Though these two aspects were discussed several other issues were also raised by participants. Three main themes emerged from the thematic analysis of the focus group data. These were Human Factors, Clinical Management and Protecting Patients. A further meta-theme was also identified as The System and Human Side of Caring.

Human Factors

The theme of Human Factors related to aspects of the environment such as design and staffing, the use of checklists and incident reporting. Aspects of physical environment were viewed as carrying potential risks and hazards to patients and the nurse is important in constantly checking equipment to ensure safety. For example, this participant stated that, *'I have to go round everywhere, checking the emergency crash call, check the monitors. The date they were serviced.'* (Critical Care group). Other participants recognised environmental design that has improved patient safety, such as, laminated flooring, '*We have a laminated grip flooring. They can still have a fall but it is much better for them.'* (Medical ward group).

The ratio of the numbers of patients to nurses was a concern, for example, *'Even in the current era, the ratio of nurses to patients is still a bit high. In terms of care, sometimes we are under so much pressure.'* (Medical Surgical group). All the groups mentioned the use of checklists. The operating theatre group mentioned the use of the World Health Organisation (WHO) checklist and the ward groups mentioned the use of intentional rounding. Though the content of these checklists are different they were seen as having advantages for patient safety and have been embedded in nursing practice. For example, '*We're very serious about protocols and policies as well and….we live by the checklist now*.' (Operating theatre group) and '*We have a checklist now and we check every single patient on the ward is safe.'* (Medical Surgical ward group).

There were ambivalent feelings regarding the use of incident reports where some participants viewed them as positive opportunities to learn from error, for example, '*You can learn from error, you can see it's not about blame culture.*' (Medical ward group), or were seen negatively, as this participant articulated, *'Yeah, a weapon, not something to help you. We're going to tell on you.*'(Critical care group). All of these approaches are systematic ways of managing error that are evident in nursing practice and the participants recognised the importance of these approaches to patient safety.

Clinical Management

The theme of Clinical Management related to communication processes and management behaviours that were relevant to the day-to-day management of patient care. Structured approaches to communication, such as handover, team briefing, and ward rounds were viewed as important for patient safety. Generally communication between nursing teams was seen as positive but communication between professions was identified as problematic, *'I think communication between nurses is good and between doctors and doctors is very good, but I think that there is a massive communication breakdown in people from different professions….I think information is lost all the time*.'(Critical care group) The role of the medical notes was viewed as being very important in communicating medical decisions to nursing staff but this was problematic for many participants. For example, *'Sometimes you are on night shift and you handover to the nurse who is taking over in the morning and you handover things that have happened and there's nothing written in the notes, nothing written*

by the doctors.' (Critical care group). The nurses perceived medical staff as not understanding the significance of the medical record for safe nursing care.

Manager behaviour was also identified as very important for the participants feeling supported in patient safety. Managers who were seen as approachable and proactive in managing patient safety were generally viewed as providing support for example, '*My manager tends to pay a lot of attention to those small details where the chart is not updated, he will remind staff, so he is very picky on the small things, which is good because it reminds everybody about what you are doing.*' (Medical Surgical group). Those managers who were seen as unsupportive tended to be reactive and not supportive of staff, for example, '*Just telling me what to do. It's just like another surgeon telling me what to do.'* (Operating theatre group).

Protecting patients

Protecting patients was a key theme that emerged as being important aspect of nursing practice relating to patient safety. This focused upon how nursing skill is applied to patient care and acting as a gatekeeper and advocate for patients. There was an overall sense that patients are vulnerable, for example, '*The nature of our patients we're receiving acutely unwell patients who are suffering from delirium and are vulnerable.*' (Medical ward group). There was a sense that nurses protect patients by ensuring safety whilst undertaking nursing tasks, for example, *'Administrating medication is a major thing and I think safety should be ensured all the time and I see we always check, because you've got a critically ill patient and the last thing you want is a drug error.'* (Critical care group). There was also a sense that nurses need to challenge others. For example, *'I think when it comes to patient safety everyone has to take responsibility for safety, the doctors just don't do it. We encourage, we try to make everyone to be attentive but you have to challenge them.'* (Critical care group). There was a clear sense that the participants felt that they had a role in protecting patients from harm.

Joint synthesis of survey and focus group findings

The results of the cross-sectional survey found a variation in the dimension of Communication between nurses working in critical care and medical wards, though pairwise

comparisons were not statistically significant. Table three shows the mean and 95% confidence intervals for the dimension of Communication and a summary of the themes identified in the thematic analysis of the focus group discussions. The ward focus groups identified nurse-to-nurse communication as important for patient safety and these groups had slightly higher safety climate scores in this area. The Critical Care and Operating theatre focus group highlighted challenges associated with nurse to doctor communication.

Table Three: Differences in the dimension of communication between critical care, operating theatres, medical and surgical wards for the SCQ and theme

SCQ Communication score	Summary of thematic analysis
Critical care Mean = 3.19 (3.04, 3,33)	The main mechanism for communication was the ward round. Problems were identified where communication was poor following a ward round or where medical staff do not record in the medical record.
Operating theatres Mean = 3.17 (2.96, 3.38)	The main mechanism for communication was the WHO checklist and team briefing. There were challenges associated with compliance with these approaches from surgeons.
Medical wards Mean = 3.50 (3.33, 3.67) Surgical wards Mean = 3.35 (3.15, 3.54)	The main focus of communication was related to handovers between nursing teams and ward rounds. These seem to work well.

There was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses, compared with nurses working in medical and surgical ward settings, with operating theatres having a lower score for Management Commitment. Table four shows the mean and confidence intervals for the dimension of Management Commitment and the themes that were identified in the focus groups. The operating theatre group reported more reactive and unsupportive manager behaviours in the focus group discussion. Whereas, the other areas generally reported proactive and supportive

manager behaviours in the focus groups, the operating theatre focus group reported reactive style of management.

Table Four: Differences in the dimension of management commitment to patient safety, between critical care, operating theatres, medical and surgical wards for the SCQ and themes.

SCQ Management commitment score	Summary of thematic analysis
Critical care Mean = 3.48 (3.34, 3.62)	Being approachable and accessible to support staff. Having experience and clinical credibility.
Operating theatres Mean = 3.27 (3.07, 3.67)	The perception that manager take sides with medical staff, not providing help and advice to nurses when they approach managers for assistance, and having an agenda related to targets, managers side with the surgeons and do not support the nursing staff, that the rules do not apply to surgeons.
Medical wards Mean = 3.75 (3.39, 3.91) Surgical wards Mean = 3.66 (3.47, 3.85)	Being proactive in supporting patient safety and reminding staff about compliance to safety procedures. Working clinically in the area and having clinical credibility with the nursing staff was highly valued and being approachable and accessible to nursing staff when they feel that they need support with problems related to patient safety.

The System and Human Side of Caring

A meta-theme, or overarching theme was identified from the three main themes and was labelled, the system and the human side of caring. This holistic view of the data captures two aspects of patient safety that seemed to be apparent within the data. That is, the system in which caring takes place, and this includes the physical environment, the design of that environment, and the system processes that have been put in place to assist patient safety with the use of checklists and incident reporting. These systematic organisational structures and processes provide the backdrop and the context in which caring takes place. The human side of caring includes the personal and the interpersonal aspects of care, the need to communicate within nursing teams and to handover care to each other. The relationship with clinical

managers was important to provide support for safe clinical care. The importance of interaction with other disciplines and the problems associated with that was a key component. Finally, the acknowledgement of patient's vulnerability within the system, and that nurses feel it is an important aspect of their role to act as an advocate and to protect patients through acting as a gatekeeper. Safety lies within an interaction between these two aspects of the clinical environment.

DISCUSSION

The application of High Reliability Organisation theory has underpinned the approach to patient safety in the past decade in the UK.[12] and the introduction of Human Factor approaches to patient safety is high on the agenda in the UK at the present time. The results of this study indicates that though human factor approaches are an important aspect of safe nursing practice, these approaches need to be supported with communication and management behaviours that rely upon good interpersonal skills. Attitudes and organisational culture are shaped and developed within the context of the transpersonal and the results of the focus groups indicate that support and communication empower nurses to advocate and protect their patients. This study highlights the importance of looking beyond systems and processes to ensure patient safety. The advent and development of checklists, the implementation of human factor and high reliability approaches are important and these have had a significant impact on patient safety but this study highlights other aspects of social behaviour and communication that can have an impact on patient safety. Indeed, too much focus upon targets and processes can be counterproductive. [34]

The SCQ has been used in the NHS extensively, however, the factor structure had not validated within a healthcare population before its use. The results of the confirmatory factor analysis undertaken here with a nursing sample, indicated that the SCQ did have an acceptable level of model fit for some criteria. However, some aspects of the tool do need careful consideration when used and further work needs to be undertaken before wider use in the healthcare context. It is important to ensure that tools developed in one context are evaluated for fit into another context.

The findings indicated there was a lower safety climate in operating theatres compared to ward areas for management commitment. Both critical care and operating theatre groups also scored lower for communication than medical ward areas, though this was close to, but not statistically different. This may seem surprising, given that in recent years there has been widespread introduction of High Reliability Organisation approaches into critical care units and operating theatres, such as the WHO checklist into operating theatres and the introduction of reliability and standardisation measures in intensive care units. [35-36]. However, these results are consistent with results from other countries and may indicate that there is a fundamental difference in safety climate in different clinical settings and it has been suggested that these differences are associated with the severity or complexity of the patient condition, high patient turnover, or the technological complexity of the care delivered.[16-18] The results of this mixed methods study may point to other factors associated with management and communication differences in these areas rather than the highly technical aspects of patient care associated with critical areas.

In a post Francis Inquiry [34] era, nursing care in particular has had increasing scrutiny of its practice, and these results indicate that there is a focus on safety in clinical practice and this is reflected in the perceptions and attitudes of the nurses who participated in this study. The factor scores of Personal priorities and need for safety, were consistently high across all groups, suggesting that for the participants, safety is an important priority in patient care for these nurses and this was reflected in the focus group discussions. The factor scores for Work environment were consistently low across all groups and the focus group discussions highlighted the availability of equipment, staffing, the resources and time available to undertake the work are important aspects of safety in nursing practice.

It is acknowledged that the results need to be considered in the light of a cross-sectional survey response rate of 57% and the fact that the number of participants in some of the focus groups was low. However, the response rate is similar to other work undertaken in the field and although there were low numbers in some focus groups robust data was generated. However, the results of this study raise some important issues relating to the underlying drivers of safety climate in nursing practice and the importance of using a mixed methodology to provide a deeper insight into the mechanisms driving safety climate in nursing practice. Using a mixed methodology enabled a much deeper investigation of potential factors driving safety climate. The utilisation of mixed methodology and a further

investigation of manager behaviours are potentially fruitful areas for further investigations in patient safety climate.

ACKNOWLEDGEMENTS

To the RCN for granting permission to use their version of the SCQ. To the participants who gave their time to complete the SCQ.

Contributorship statement

Maggie Tarling made a substantial contribution to the initial concept and design, data collection and analysis, drafting and revising the work and was main editior for this submission.

Anne Jones made a substantial contribution to the development of the design, qualitative analysis, dafting and revision of the work and final approval of the published version.

Trevor Murrells made a substantial contribution to the development of the design, analysis of the quantitative data, drafting and revision of the work and final approval of the published version.

Helen McCutcheon made a substantial contribution to the development of the design, qualitative data analysis, drafting and revising the work and final approval of the publilshed version.

Competing interests

There are no competing interests

Funding

The research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data sharing statement

Original data is held by the corresponding author, but arrangements are being made for data to be deposited into Dryad.

REFERENCES

1. The Health Foundation. Measuring Safety Culture. London: The Health Foundation 2011.
2. The Health Foundation. The measurement and monitoring of safety. London: The Health Foundation. 2013.
3. Health and Safety Commission (HSC). Organising for Safety: Third Report. ACSNI Study Group on Human Factors. Sudbury: HSE Books. 1993 P15.
4. Health and Safety Executive. Reducing error and influencing behaviour. London: HSE. 1999.
5. Pronovost P and Sexton B. Assessing safety culture: guidelines and recommendations. *Qual Saf Health Care*. 2005;14:231–233.
6. Freeth D, Sandall J, Allan T, et al. A methodological study to compare survey-based and observation-based evaluations of organisational and safety cultures and then compare both approaches with makers of the quality of care. Health Technology Assessment. 2012;6:25.
7. Cox SJ and Cheyne AJT. Assessing safety culture in offshore environments. *Saf Sci.* 2000;34:111–129.
8. Currie L and Watterson L. Measuring the safety climate in NHS organisations. *Nurs Stand*. 2010;24(24):35–38.
9. University of Loughborough. Safety climate measurement user guide and toolkit. 2000. Available at: http://www.lboro.ac.uk/departments/sbe/downloads/pmdc/safety-climate-assessment-toolkit.pdf (assessed July 2013)
10. The Health Foundation. High Reliability Organisations. London: The Health Foundation. 2011.
11. Health and Safety Executive. High reliability organisations: A review of the literature. London: HSE. 2011.
12. Department of Health. An Organisation with a memory. London: The Stationary Office. 2001.
13. Olsen E and Aase KA. Comparative study of safety climate differences in healthcare and the petroleum industry. *Qual Saf Health Care.* 2010;19(Suppl 3):175–179.
14. Currie L and Cooper C. Measuring staffs' perceptions of safety climate: testing The Safety Climate Questionnaire (SCQ) for use in healthcare. London: RCN. 2008.
15. Zohar D, Livne Y, Tenne-Gazit O, et al. Healthcare climate: a framework for measuring and improving patient safety. *Critical Care Medicine.* 2007;35:1312–1317.

16. Hartmann CW, Rosen AK, Meterfo M, et al. An overview of patient safety climate in the VA. *Health Services Research.* 2008;43(4):1263–1284.
17. Singer SJ, Gaba DM, Fulwell A, et al. Patient safety climate in 92 US hospital: differences by work area and discipline. *Medical Care.* 2009;47:23–31.
18. Campbell EG, Singer S, Kitch BT, et al. Patient Safety Climate in Hospitals: Act Locally on Variation Across Units. *Joint Commission on Accreditation of Healthcare Organizations.* 2010;36(7):316–326.
19. Abbas HAE, Bassiuni NA, Baddar FM Perception of front-line healthcare providers toward patient safety: A preliminary study in a university hospital in Egypt. Topics in Advanced Practice Nursing e-journal. 2008:8(2): Medscape.
20. Abdou HA and Saber KM A baseline assessment of patient safety culture among nurses at student university hospital. World Journal of Medical Sciences. 2011:6(1): 17 – 26.
21. Needleman J, Buerhaus P, Pankratz VS, et al. Nurse staffing and inpatient hospital mortality. *N Engl J Med.* 2011;364:1037-1045.
22. Aiken LH, Sermeus W, Van den Heede K et al Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. British Medical Journal. 2012:344 e1717.
23. Leech N and Onwuegbuzie AJ (2009) A typology of mixed methods research designs. Quality and Quantity. 43: 265 – 275
24. Glynos J and Howarth D (2007) Logics of Critical Explanation in Social and Political Theory. London: Routledge.
25. Guetterman TC, Fetters MD, Creswell JW Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med. 2015 Nov; 13(6): 554–561.
26. Comrey, AL and Lee, HB*A. First Course in Factor Analysis.* Hillsdale, NJ: Erlbaum. 1992.
27. Elliot AC, Hynan LS, Reisch S, et al. Preparing data for analysis using Microsoft Excel. *J Investig Med.* 2006;54:6:334–341.
28. Little RJA and Rubin DB. Statistical analysis with missing data (2nd ed.). New York: Wiley. 2002.
29. Kline P. An easy guide to factor analysis. London: Routledge. 2008.
30. Hu L and Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*.1999;6(1):1-55.

31. Steiger JH. Understanding the limitations of global fit assessment in structural equation modelling. *Individual Differences.* 2007;42:893–898.
32. Lui RY. Bootstrap procedures under some Non-I.I.D. models. The Annals of Statistics. 1988;4:1696-1708.
33. Braun V and Clarke V Using thematic analysis in psychology. Qualitative Research in Psychology. 2006.3 (2): 77-101.
34. Francis R Report of the Mid-Staffordshire NHS Foundation Trust public inquiry. 2013 London: The Stationary Office.
35. Haynes AB, Weiser TG, Berry WR, et al. (2009) A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;360:491-499.
36. Pronovost PJ, Berenholtz SM, Goeschel CA. Creating High Reliability in Health Care Organizations. *Health Services Research.* 2006;41(4) Part II:1599–1617.

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Management Commitment			
Managers and supervisors express concern if safety procedures are not followed	0.440	0.633	0.696
In my workplace managers/supervisors show interest in my safety	0.520	0.689	0.727
In my workplace management turn a blind eye to safety issues	0.737	0.618	0.712
In my workplace management acts quickly to correct safety problems	0.811	0.743	0.806
Corrective action is always taken when management is told about unsafe practice	0.690	0.527	0.571
Management acts only after incidents have occurred	0.500	0.666	0.710
Management acts decisively when a safety concern is raised	0.792	0.668	0.732
Priority of Safety			
Management considers safety to be equally important as getting the work done	0.534	0.717	0.757
Safety rules and procedures are carefully followed	0.585	0.592	0.670
Management clearly considers the safety of staff of great importance	0.665	0.659	0.688
I believe that safety issues are not assigned a high priority	0.585	0.675	0.742
Communication			
There is good communication here about safety issues which affect me	0.731	0.596	0.653
Safety information is always brought to my attention by my line manager/supervisor	0.633	0.596	0.638
I do not receive praise for working safely	0.481	0.511	0.569
My line manager/supervisor does not always inform me of current concern and issues	0.594	0.636	0.703
Management operates an open door policy on safety issues	0.541	0.511	0.574
Safety rules			
Some safety rules and procedures do not need to be followed to get the job done safely	0.724	0.600	0.580
Some health and safety rules and procedures are not really practical	0.685	0.622	0.744
Sometimes it is necessary to depart from safety requirements in order to get the work done	0.583	0.699	0.709
Supportive environment			
I can influence health and safety here	0.543	0.460	0.479
A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate	0.367	0.150*	0.153*
I am strongly encouraged to report unsafe conditions	0.639	0.692	0.756
When people ignore safety procedures here I feel it is none of my business	0.480	0.291*	0.387*
Co-workers often give tips to each other on how to work safely	0.323	0.452	0.498
Employees are not encouraged to raise safety concerns	0.421	0.508	0.624

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Personal priorities and need for safety			
A safe place to work has a lot of personal meaning to me	0.571	0.519	0.600
It is important to me that there is a continuing emphasis on safety	0.655	0.560	0.625
I understand the safety rules for my job	0.642	0.664	0.832
Personally I feel that safety issues are not the most important aspect of my job	0.500	0.571	0.694
Safety is the number one priority in my mind when completing a job	0.623	0.617	0.685
Personal appreciation of risk			
I am clear about what my responsibilities are for health and safety	0.273	0.561	0.744
I am sure it is only a matter of time before I am involved in an incident	0.782	0.382	0.383
In my workplace the chance of being involved in an incident are high	0.464	0.548	0.570
I am rarely worried about being injured on the job	0.286	0.110*	0.086*
Involvement			
I am involved with safety issues at work	0.687	0.657	0.602
I am never involved in the ongoing review of safety	0.524	0.402	0.541
I am involved in informing management of important safety issues	0.724	0.671	0.677
Work environment			
There are always enough people available to get the job done safely	0.596	0.527	0.593
Sometimes I am not given enough time to get the job done safely	0.668	0.610	0.543
Sometimes conditions here hinder my ability to work safely	0.666	0.724	0.727
Operational targets often conflict with safety measures	0.795	0.590	0.662
I cannot always get the equipment I need to do the job safely	0.448	0.587	0.645
This is a safer place to work than other Trusts I have worked for	0.256	0.270*	0.392*

McGill

Mixed Methods Appraisal Tool (MMAT) – Version 2011

For dissemination, application, and feedback: Please contact pierre.pluye@mcgill.ca, Department of Family Medicine, McGill University, Canada.

The MMAT is comprised of two parts (see below): criteria (Part I) and tutorial (Part II). While the content validity and the reliability of the pilot version of the MMAT have been examined, this critical appraisal tool is still in development. Thus, the MMAT must be used with caution, and users' feedback is appreciated. Cite the present version as follows.

> Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'Cathain, A., Griffiths, F., Boardman, F., Gagnon, M.P., & Rousseau, M.C. (2011). *Proposal: A mixed methods appraisal tool for systematic mixed studies reviews*. Retrieved on [date] from http://mixedmethodsappraisaltoolpublic.pbworks.com. Archived by WebCite® at http://www.webcitation.org/5tTRTc9yJ

Purpose: The MMAT has been designed for the appraisal stage of complex systematic literature reviews that include qualitative, quantitative and mixed methods studies (mixed studies reviews). The MMAT permits to concomitantly appraise and describe the methodological quality for three methodological domains: mixed, qualitative and quantitative (subdivided into three sub-domains: randomized controlled, non-randomized, and descriptive). Therefore, using the MMAT requires experience or training in these domains. E.g., MMAT users may be helped by a colleague with specific expertise when needed. The MMAT allows the appraisal of most common types of study methodology and design. For appraising a qualitative study, use section 1 of the MMAT. For a quantitative study, use section 2 or 3 or 4, for randomized controlled, non-randomized, and descriptive studies, respectively. For a mixed methods study, use section 1 for appraising the qualitative component, the appropriate section for the quantitative component (2 or 3 or 4), and section 5 for the mixed methods component. For each relevant study selected for a systematic mixed studies review, the methodological quality can then be described using the corresponding criteria. This may lead to exclude studies with lowest quality from the synthesis, or to consider the quality of studies for contrasting their results (e.g., low quality vs. high).

Scoring metrics: For each retained study, an overall quality score may be not informative (in comparison to a descriptive summary using MMAT criteria), but might be calculated using the MMAT. Since there are only a few criteria for each domain, the score can be presented using descriptors such as *, **, ***, and ****. For qualitative and quantitative studies, this score can be the number of criteria met divided by four (scores varying from 25% (*) -one criterion met- to 100% (****) -all criteria met-). For mixed methods research studies, the premise is that the overall quality of a combination cannot exceed the quality of its weakest component. Thus, the overall quality score is the lowest score of the study components. The score is 25% (*) when *QUAL=1 or QUAN=1 or MM=0*; it is 50% (**) when *QUAL=2 or QUAN=2 or MM=1*; it is 75% (***) when *QUAL=3 or QUAN=3 or MM=2*; and it is 100% (****) when *QUAL=4 and QUAN=4 and MM=3* (QUAL being the score of the qualitative component; QUAN the score of the quantitative component; and MM the score of the mixed methods component).

Rationale: There are general criteria for planning, designing and reporting mixed methods research (Creswell and Plano Clark, 2010), but there is no consensus on key specific criteria for appraising the methodological quality of mixed methods studies (O'Cathain, Murphy and Nicholl, 2008). Based on a critical examination of 17 health-related systematic mixed studies reviews, an initial 15-criteria version of MMAT was proposed (Pluye, Gagnon, Griffiths and Johnson-Lafleur, 2009). This was pilot tested in 2009. Two raters assessed 29 studies using the pilot MMAT criteria and tutorial (Pace, Pluye, Bartlett, Macaulay et al., 2010). Based on this pilot exercise, it is anticipated that applying MMAT may take on average 15 minutes per study (hence efficient), and that the Intra-Class Correlation might be around 0.8 (hence reliable). The present 2011 revision is based on feedback from four workshops, and a comprehensive framework for assessing the quality of mixed methods research (O'Cathain, 2010).

Conclusion: The MMAT has been designed to appraise the *methodological quality* of the studies retained for a systematic mixed studies review, not the quality of their *reporting* (writing). This distinction is important, as good research may not be 'well' reported. If reviewers want to genuinely assess the former, companion papers and research reports should be collected when some criteria are not met, and authors of the corresponding publications should be contacted for additional information. Collecting additional data is usually necessary to appraise *qualitative research and mixed methods studies*, as there are no uniform standards for reporting study characteristics in these domains (www.equator-network.org), in contrast, e.g., to the CONSORT statement for reporting randomized controlled trials (www.consort-statement.org).

Authors and contributors: Pierre Pluye[1], Marie-Pierre Gagnon[2], Frances Griffiths[3] and Janique Johnson-Lafleur[1] proposed an initial version of MMAT criteria (Pluye et al., 2009). Romina Pace[1] and Pierre Pluye[1] led the pilot test. Gillian Bartlett[1], Belinda Nicolau[4], Robbyn Seller[1], Justin Jagosh[1], Jon Salsberg[1] and Ann Macaulay[1] contributed to the pilot work (Pace et al., 2010). Pierre Pluye[1], Émilie Robert[5], Margaret Cargo[6], Alicia O'Cathain[7], Frances Griffiths[3], Felicity Boardman[3], Marie-Pierre Gagnon[2], Gillian Bartlett[1], and Marie-Claude Rousseau[8] contributed to the present 2011 version.

Affiliations: 1. Department of Family Medicine, McGill University, Canada; 2. Faculté des sciences infirmières, Université Laval, Canada; 3. Warwick Medical School, University of Warwick, UK; 4. Faculty of Dentistry, McGill University, Canada; 5. Centre de recherche du CHUM, Université de Montréal, Canada; 6. School of Health Sciences, University of South Australia, Australia; 7. Medical Care Research Unit, ScHARR, University of Sheffield, UK; 8. INRS-Institut Armand Frappier, Laval, Canada.

PART I. MMAT criteria & one-page template (to be included in appraisal forms)

Types of mixed methods study components or primary studies	Methodological quality criteria (see tutorial for definitions and examples)	Responses			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	• Are there clear qualitative and quantitative research questions (or objectives*), or a clear mixed methods question (or objective*)?				Page 5
	• Do the collected data allow address the research question (objective)? E.g., consider whether the follow-up period is long enough for the outcome to occur (for longitudinal studies or study components).				Page 5
	Further appraisal may be not feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.				
1. Qualitative	1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?				Page 8
	1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?				Page 8
	1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?				Page 8 & Page 18
	1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?				Page 8
2. Quantitative randomized controlled (trials)	2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?				N/A
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)?				N/A
	2.3. Are there complete outcome data (80% or above)?				N/A
	2.4. Is there low withdrawal/drop-out (below 20%)?				N/A
3. Quantitative non-randomized	3.1. Are participants (organizations) recruited in a way that minimizes selection bias?				N/A
	3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?				N/A
	3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?				N/A
	3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?				N/A
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?				Page 6
	4.2. Is the sample representative of the population understudy?				Page 6
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?				Page 6
	4.4. Is there an acceptable response rate (60% or above)?				Page 8 & page 18
5. Mixed methods	5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?				Page 6
	5.2. Is the integration of qualitative and quantitative data (or results*) relevant to address the research question (objective)?				Page 14 - 16
	5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results*) in a triangulation design?				Page 18
	Criteria for the qualitative component (1.1 to 1.4), and appropriate criteria for the quantitative component (2.1 to 2.4, or 3.1 to 3.4, or 4.1 to 4.4), must be also applied.				

*These two items are not considered as double-barreled items since in mixed methods research, (1) there may be research questions (quantitative research) or research objectives (qualitative research), and (2) data may be integrated, and/or qualitative findings and quantitative results can be integrated.

PART II. MMAT tutorial

Types of mixed methods study components or primary studies	Methodological quality criteria
1. Qualitative Common types of qualitative research methodology include: A. Ethnography The aim of the study is to describe and interpret the shared cultural behaviour of a group of individuals. B. Phenomenology The study focuses on the subjective experiences and interpretations of a phenomenon encountered by individuals. C. Narrative The study analyzes life experiences of an individual or a group. D. Grounded theory Generation of theory from data in the process of conducting research (data collection occurs first). E. Case study In-depth exploration and/or explanation of issues intrinsic to a particular case. A case can be anything from a decision-making process, to a person, an organization, or a country. F. Qualitative description There is no specific methodology, but a qualitative data collection and analysis, e.g., in-depth interviews or focus groups, and hybrid thematic analysis (inductive and deductive). Key references: Creswell, 1998; Schwandt, 2001; Sandelowski, 2010.	**1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?** E.g., consider whether (a) the selection of the participants is clear, and appropriate to collect relevant and rich data; and (b) reasons why certain potential participants chose not to participate are explained. **1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?** E.g., consider whether (a) the method of data collection is clear (in depth interviews and/or group interviews, and/or observations and/or documentary sources); (b) the form of the data is clear (tape recording, video material, and/or field notes for instance); (c) changes are explained when methods are altered during the study; and (d) the qualitative data analysis addresses the question. **1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?*** E.g., consider whether the study context and how findings relate to the context or characteristics of the context are explained (how findings are influenced by or influence the context). "For example, a researcher wishing to observe care in an acute hospital around the clock may not be able to study more than one hospital. (…) Here, it is essential to take care to describe the context and particulars of the case [the hospital] and to flag up for the reader the similarities and differences between the case and other settings of the same type" (Mays & Pope, 1995). The notion of context may be conceived in different ways depending on the approach (methodology) tradition. **1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?*** E.g., consider whether (a) researchers critically explain how findings relate to their perspective, role, and interactions with participants (how the research process is influenced by or influences the researcher); (b) researcher's role is influential at all stages (formulation of a research question, data collection, data analysis and interpretation of findings); and (c) researchers explain their reaction to critical events that occurred during the study. The notion of reflexivity may be conceived in different ways depending on the approach (methodology) tradition. E.g., "at a minimum, researchers employing a generic approach [qualitative description] must explicitly identify their disciplinary affiliation, what brought them to the question, and the assumptions they make about the topic of interest" (Caelli, Ray & Mill, 2003, p. 5).

*See suggestion on the MMAT wiki homepage (under '2011 version'): Independent reviewers can establish a common understanding of these two items prior to beginning the critical appraisal.

Types of mixed methods study components or primary studies	Methodological quality criteria
2. Quantitative randomized controlled (trials) Randomized controlled clinical trial: A clinical study in which individual participants are allocated to intervention or control groups by randomization (intervention assigned by researchers). Key references: Higgins & Green, 2008; Porta, 2008; Oxford Center for Evidence based medicine, 2009.	**2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?** In a randomized controlled trial, the allocation of a participant (or a data collection unit, e.g., a school) into the intervention or control group is based solely on chance, and researchers describe how the randomization schedule is generated. "A simple statement such as 'we randomly allocated' or 'using a randomized design' is insufficient". *Simple randomization:* Allocation of participants to groups by chance by following a predetermined plan/sequence. "Usually it is achieved by referring to a published list of random numbers, or to a list of random assignments generated by a computer". *Sequence generation:* "The rule for allocating interventions to participants must be specified, based on some chance (random) process". Researchers provide sufficient detail to allow a readers' appraisal of whether it produces comparable groups. E.g., blocked randomization (to ensure particular allocation ratios to the intervention groups), or stratified randomization (randomization performed separately within strata), or minimization (to make small groups closely similar with respect to several characteristics).
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)? *The allocation concealment protects assignment sequence until allocation.* E.g., researchers and participants are unaware of the assignment sequence up to the point of allocation. E.g., group assignment is concealed in opaque envelops until allocation. *The blinding protects assignment sequence after allocation.* E.g., researchers and/or participants are unaware of the group a participant is allocated to during the course of the study.
	2.3. Are there complete outcome data (80% or above)? E.g., almost all the participants contributed to almost all measures.
	2.4. Is there low withdrawal/drop-out (below 20%)? E.g., almost all the participants completed the study.

Types of mixed methods study components or primary studies	Methodological quality criteria
3. Quantitative non-randomized Common types of design include (A) non-randomized controlled trials, and (B-C-D) observational analytic study or component where the intervention/exposure is defined/assessed, but not assigned by researchers. A. Non-randomized controlled trials The intervention is assigned by researchers, but there is no randomization, e.g., a pseudo-randomization. A non-random method of allocation is not reliable in producing alone similar groups. B. Cohort study Subsets of a defined population are assessed as exposed, not exposed, or exposed at different degrees to factors of interest. Participants are followed over time to determine if an outcome occurs (prospective longitudinal). C. Case-control study Cases, e.g., patients, associated with a certain outcome are selected, alongside a corresponding group of controls. Data is collected on whether cases and controls were exposed to the factor under study (retrospective). D. Cross-sectional analytic study At one particular time, the relationship between health-related characteristics (outcome) and other factors (intervention/exposure) is examined. E.g., the frequency of outcomes is compared in different population sub-groups according to the presence/absence (or level) of the intervention/exposure. Key references for observational analytic studies: Higgins & Green, 2008; Wells, Shea, O'Connell, Peterson, et al., 2009.	**3.1. Are participants (organizations) recruited in a way that minimizes selection bias?** At recruitment stage: For cohort studies, e.g., consider whether the exposed (or with intervention) and non-exposed (or without intervention) groups are recruited from the same population. For case-control studies, e.g., consider whether same inclusion and exclusion criteria were applied to cases and controls, and whether recruitment was done independently of the intervention or exposure status. For cross-sectional analytic studies, e.g., consider whether the sample is representative of the population. **3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?** At data collection stage: E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) the measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure. For non-randomized controlled trials, the intervention is assigned by researchers, and so consider whether there was absence/presence of a contamination. E.g., the control group may be indirectly exposed to the intervention through family or community relationships. **3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** At data analysis stage: For cohort, case-control and cross-sectional, e.g., consider whether (a) the most important factors are taken into account in the analysis; (b) a table lists key demographic information comparing both groups, and there are no obvious dissimilarities between groups that may account for any differences in outcomes, or dissimilarities are taken into account in the analysis. **3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?**

Types of mixed methods study components or primary studies	Methodological quality criteria
4. Quantitative descriptive studies Common types of design include single-group studies: A. Incidence or prevalence study without comparison group In a defined population at one particular time, what is happening in a population, e.g., frequencies of factors (importance of problems), is described (portrayed). B. Case series A collection of individuals with similar characteristics are used to describe an outcome. C. Case report An individual or a group with a unique/unusual outcome is described in details. Key references: Critical Appraisal Skills Programme, 2009; Draugalis, Coons & Plaza, 2008.	**4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?** E.g., consider whether (a) the source of sample is relevant to the population under study; (b) when appropriate, there is a standard procedure for sampling, and the sample size is justified (using power calculation for instance).
	4.2. Is the sample representative of the population understudy? E.g., consider whether (a) inclusion and exclusion criteria are explained; and (b) reasons why certain eligible individuals chose not to participate are explained.
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)? E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure.
	4.4. Is there an acceptable response rate (60% or above)? The response rate is not pertinent for case series and case report. E.g., there is no expectation that a case series would include all patients in a similar situation.

Types of mixed methods study components or primary studies	Methodological quality criteria
5. Mixed methods Common types of design include: A. Sequential explanatory design The quantitative component is followed by the qualitative. The purpose is to explain quantitative results using qualitative findings. E.g., the quantitative results guide the selection of qualitative data sources and data collection, and the qualitative findings contribute to the interpretation of quantitative results. B. Sequential exploratory design The qualitative component is followed by the quantitative. The purpose is to explore, develop and test an instrument (or taxonomy), or a conceptual framework (or theoretical model). E.g., the qualitative findings inform the quantitative data collection, and the quantitative results allow a generalization of the qualitative findings. C. Triangulation design The qualitative and quantitative components are concomitant. The purpose is to examine the same phenomenon by interpreting qualitative and quantitative results (bringing data analysis together at the interpretation stage), or by integrating qualitative and quantitative datasets (e.g., data on same cases), or by transforming data (e.g., quantization of qualitative data). D. Embedded design The qualitative and quantitative components are concomitant. The purpose is to support a qualitative study with a quantitative sub-study (measures), or to better understand a specific issue of a quantitative study using a qualitative sub-study, e.g., the efficacy or the implementation of an intervention based on the views of participants. Key references: Creswell & Plano Clark, 2007; O'Cathain, 2010.	**5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?** E.g., the rationale for integrating qualitative and quantitative methods to answer the research question is explained. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question (objective)?** E.g., there is evidence that data gathered by both research methods was brought together to form a complete picture, and answer the research question; authors explain when integration occurred (during the data collection-analysis or/and during the interpretation of qualitative and quantitative results); they explain how integration occurred and who participated in this integration. **5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results)?**

References

- Caelli, K., Ray, L., & Mill, J. (2003). 'Clear as Mud': Toward greater clarity in generic qualitative research. *International Journal of Qualitative Methods, 2(*2), 1-23.
- Creswell, J., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. London: Sage.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks: Sage.
- Critical Appraisal Skills Programme (2009). CASP appraisal tools. Retrieved on August 26, 2009 from: *www.phru.nhs.uk/pages/PHD/resources.htm*
- Draugalis, J.R., Coons, S.J., & Plaza, C.M. (2008). Best practices for survey research reports: a synopsis for authors and reviewers. *American Journal of Pharmaceutical Education, 72(1),* e11.
- Higgins, J.P.T. & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* - Version 5.0.1 [updated September 2008]. The Cochrane Collaboration. Retrieved on August 26, 2009 from www.cochrane-handbook.org
- Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *British Medical Journal, 311(6997),* 109-112.
- O'Cathain, A., Murphy, E. & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research and Policy, 13(2),* 92-98.
- O'Cathain, A. (2010). Assessing the quality of mixed methods research: Towards a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research (2nd edition)* (pp. 531-555). Thousand Oaks: Sage.
- Pace, R., Pluye, P., Bartlett, G., Macaulay, A., Salsberg, J., Jagosh, J., & Seller, R. (2010). *Reliability of a tool for concomitantly appraising the methodological quality of qualitative, quantitative and mixed methods research: a pilot study*. 38th Annual Meeting of the North American Primary Care Research Group (NAPCRG), Seattle, USA.
- Pluye, P., Gagnon, M.P., Griffiths, F. & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46(4),* 529-46.
- Oxford Center for Evidence Based Medicine (2009). Levels of evidence. Retrieved on July 7, 2009 from www.cebm.net/levels_of_evidence.asp
- Porta, M. (2008). *A Dictionary of Epidemiology*. New York: Oxford University Press.
- Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Research in Nursing and Health, 33*(1), 77-84.
- Schwandt, T. (2001). *Dictionary of qualitive inquiry*. Thousand Oaks: Sage.
- Wells, G.A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2009). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses*. The Cochrane Non-Randomized Studies Method Group. Retrieved on July 7, 2009 from www.ohri.ca/programs/clinical_epidemiology/oxford.htm

BMJ Open

Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016977.R1
Article Type:	Research
Date Submitted by the Author:	13-Jun-2017
Complete List of Authors:	Tarling, Maggie; City University School of Health Sciences, Nursing Jones, Anne; Kings College London, Adult Nursing Murrells, Trevor; Kings College London, National Nursing Research Unit McCutcheon, Helen; University of Queensland, Faculty of Health and Behavioural Sciences
Primary Subject Heading:	Nursing
Secondary Subject Heading:	Communication, Health services research, Intensive care, Qualitative research
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Adult intensive & critical care < INTENSIVE & CRITICAL CARE, QUALITATIVE RESEARCH

Title: Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Keywords: patient safety; safety climate; nursing; mixed methods; human factors

Word count: 5057

ABSTRACT

Objectives: To explore the underlying factors that influence the presence of safety climate in nursing practice.

Design: A sequential mixed methods design included a cross-sectional survey using the Safety Climate Questionnaire (SCQ) and thematic analysis of focus group discussions. Confirmatory Factor Analysis (CFA) was used to validate the factor structure of the SCQ. Factor scores were compared between nurses working in operating theatres, critical care and ward areas. Results from the survey and the thematic analysis were then compared and synthesised.

Setting: A London University.

Participants: 319 registered nurses working in acute hospital settings completed the SCQ and a further 23 nurses participated in focus groups.

Results: CFA indicated that there was a good model fit on some criteria (χ^2 = 1683.699, df 824, $p < 0.001$; χ^2/ df = 2.04; RMSEA = 0.058) but a less acceptable fit on Comparative Fit Index (CFI) = 0.804. There was a statistically significant difference between clinical specialisms in Management Commitment ($F [4,266] = 4.66$, $p = 0.001$). Nurses working in operating theatres had lower scores compared with ward areas and they also reported negative perceptions about management in their focus group. There was significant variation in scores for Communication across clinical specialism ($F [4,266] = 2.62$, $p = 0.035$) but none of the pair-wise comparisons achieved statistical significance. Thematic analysis identified themes of Human Factors, Clinical Management and Protecting Patients. The System and the Human Side of Caring was identified as a meta-theme. There were areas of overlap and differences between the dimensions of the SCQ and the findings of the thematic analysis.

Conclusions: The results suggest that the SCQ has some utility but requires further exploration. The findings indicate that safety in nursing practice is a complex interaction between safety systems and the social and interpersonal aspects of clinical practice.

ARTICLE SUMMARY

Strengths and limitations of this study:

- The results of the study indicate that there is an important and complex link between human factor approaches used in nursing practice and the interpersonal aspects of care.
- This work makes a unique contribution to understanding safety climate in nursing practice in the UK setting.
- The Confirmatory Factor Analysis of the Safety Climate Questionnaire indicated that the model fit could be improved but further psychometric exploratory analysis may be warranted.
- The results need to be considered in the light of a cross-sectional survey response rate of 57% and a low number of participants in some of the focus groups.

The research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Original data is held by the corresponding author but arrangements are being made for data to be deposited onto Dryad.

Author contribution statement:

M Tarling made a substantial contribution to the initial concept and design, data collection and analysis, drafting and revising the work and was main editor for this submission.

A Jones made a substantial contribution to the development of the design, qualitative analysis, drafting and revision of the work and final approval of the published version.

T Murrells made a substantial contribution to the development of the design, analysis of the quantitative data, drafting and revision of the work and final approval of the published version.

H McCutcheon made a substantial contribution to the development of the design, qualitative data analysis, drafting and revising the work and final approval of the published version.

INTRODUCTION

There is a growing consensus in healthcare safety research that organisational culture is critical for patient safety [1] and that safety management should move away from depending on lagging indicators of safety issues, such as incident reports, and move towards leading indicators, such as, measures of safety climate.[2] Patient safety culture is defined as aspects of organisational culture that are 'the product of individual and group values, attitudes, perceptions, competencies and patterns of behaviour that determine the commitment to and the style and proficiency of an organisations' health and safety management'.[3] Safety climate is defined as a measurable feature of staff's attitudes and perceptions of an organisations underlying safety culture at any point in time.[4] There is evidence that safety climate is open to change and has an impact on individual safety behaviour and an important factor in improving patient safety.[5,6]

The Safety Climate Questionnaire (SCQ) developed in the UK [7] has been used extensively in the NHS by the Royal College of Nursing.[8] However the SCQ was originally developed for use in the UK petroleum industry as part of a tool kit to measure safety climate. The SCQ measures nine factors that contribute to safety climate, namely Management commitment, Communication, Priority of safety, Safety rules and Procedures, Supportive environment, Involvement, Personal priorities and need for safety, Personal appreciation of risk and Work environment.[9] It is noted that the petroleum industry exhibits aspects of a High Reliability Organisation,[10] defined as 'organisations that are able to manage and sustain almost error-free performance despite operating in hazardous conditions where the consequences of errors could be catastrophic'[11] and as such lessons learnt from High Reliability Organisations have underpinned developments in safety and risk management in the NHS.[12] The petroleum industry is a very different setting from healthcare organisations but it is possible that their safety management systems could provide beneficial outcomes in safety and risk management in the healthcare setting.[13] A pilot of the SCQ undertaken within the NHS tested its usability and found that the tool was useable in this context. [14] However, neither an exploratory or confirmatory factor analysis of the tool was undertaken to validate its psychometric properties with a healthcare population.

Research evidence suggests that measures of safety climate vary between and within healthcare organisations and that there is limited understanding of the factors that may

influence and explain the sources of these variations.[15] Several research studies have reported safety climate scores varying across different clinical specialities with some reporting less safe climates in operating theatres, critical care and emergency departments compared to surgical and medical inpatient areas [16-18] and others reporting a safer climate in critical care [19-20]. However, none of this research has been undertaken in the UK. The underlying reasons for these variations in safety climate are unclear at the present time.

As a subset of healthcare practitioners, nurses make an important contribution to patient care and evidence indicates that nurse-staffing levels have a direct impact on patient mortality [21-22] and nurse's perceptions of safety climate levels have a direct impact on safety behaviours and outcomes. [23] Therefore it is important to understand how nurses perceive safety climate as this may have a direct impact on patient safety. This mixed methods study set out to explore the underlying factors that contribute to safety climate in nursing practice. The main aim of the study was to explore the potential sources of variation in safety climate that have been found between different clinical specialities. The study set out to determine whether there are differences in the perception of safety climate between nurses working in critical care, operating theatres, surgical and medical wards in acute hospital settings in the UK and understand the meaning that nurses working in these different clinical settings attribute to their understanding of patient safety. The factor structure of the SCQ was also explored.

METHOD

The study design was a fully mixed, sequential, equal status, mixed methods design and was conducted in two phases.[24] The first phase of the study measured and then compared safety climate scores between groups of nurses working in operating theatres, critical care, surgical and medical ward areas. As the factor structure of the SCQ had not been evaluated in a nursing sample a Confirmatory Factor Analysis (CFA) was also undertaken. The results from the cross-sectional survey were used to structure the focus group discussions held with groups of nurses from operating theatres, critical care and ward areas. The results of both phases of the study were then jointly summarised in a statistics-by-theme format to facilitate more in-depth inferences in order to consider potential mechanisms underlying safety climate.[25-26]

Following local ethical approval participants were recruited from a qualified nursing population who attended a university that recruited from a wide range of NHS Trusts and private hospitals in the region. In the UK Band 5 and 6 nurses are qualified nurses who deliver bedside care. They were specifically chosen, because they have a direct impact upon patient care and safety in their everyday practice. A convenience sampling method was used and participants were approached by the researcher at the beginning of a teaching session and the purpose of the survey was explained. Information sheets were included with the questionnaire and completion of the questionnaire implied consent. All questionnaires distributed were collected at the end of the afternoon teaching session. The aim was to collect at least 300 questionnaires as this is considered by some to be the minimum number required for robust factor analysis.[27]

A paper version of the SCQ was distributed to participants. Additional questions were added to the questionnaire in order to facilitate a stratified analysis to compare scores between nurses working in different clinical settings and measure potential factors that may influence perceptions of safety. These additional questions collected data on the clinical area the participant worked in, including whether they worked in a surgical ward, medical ward, critical care unit, operating theatre or other acute hospital unit. Further information included how long they had worked in their present position, how long they had worked in the speciality, how long they had been qualified and whether they had safety training and further training in their speciality. Participants were also asked to describe the type of training they had undertaken.

The SCQ has 43 questions with a 5-point Likert scale response and is scored by allocating a value of 5 to the ‘strongly agree’ response, 4 to ‘agree’ response, 3 to the ‘neither agree nor disagree’ response, 2 to the ‘disagree’ response and 1 to the ‘strongly disagree’ response. The negative worded questions were allocated a reverse score by subtracting the initial score from 6. The initial scores from the questionnaires provided raw scores and these were transferred into an Excel® 2013 spreadsheet. In order to ensure that the data entry was as accurate as possible a double data entry procedure was followed as recommended by Elliot et al.[28].

The Excel spreadsheet was then transferred into SPSS® V21 and a Little's 'missing completely at random' (MCAR) test was undertaken to ensure that any missing data was not introducing bias into the analysis.[29]

A confirmatory factor analysis of the SCQ scores was undertaken using SPSS® Amos V21. The original nine factor structure as identified by Cox and Cheyne was used as the *a priori* model to be confirmed by the factor analysis.[7] The following goodness of fit indices were used to test the model. Chi Square (χ^2) and the χ^2/df ratio, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The χ^2/df ratio overcomes the problem of a statistically significant χ^2 result associated with a larger sample sizes. A value of between 2 -3 is deemed as being acceptable the smaller value the better the fit.[30] The Comparative Fit Index (CFI) measures the difference in the non-centrality estimates of the baseline and proposed model with values ranging from 0 to 1. A cutoff value above 0.9 is considered to be an indication of a good model fit. The Root Mean Square Error of Approximation (RMSEA) measures the discrepancy between the hypothesised model and the population covariance matrices, and values range from 0 – 1. A RMSEA of less than 0.06 is indicative of an acceptable model fit with a recommended upper limit of 0.07.[31-32]

Once the CFA had been undertaken comparisons of safety climate dimensions (factors) were made between different clinical settings. Higher mean scores indicate a good safety climate. Dimension scores were compared between clinical specialisms using a general linear model (GLM) that adjusted for the following characteristics: years in current position, years qualified, years in specialism, specialist qualification and safety training. Adjusted means with 95%confidence intervals were calculated. Where there were differences between clinical specialism, based on the GLM F statistic, Bonferroni post-hoc pair wise comparisons were performed.

A Levene test of homogeneity of variance was conducted and residual plots produced, to ascertain whether the assumptions underpinning GLM had been met. A wild Bootstrap analysis was undertaken on the 'Personal priorities and need for safety' dimension to assess whether non-equality of variance had biased the results. [33] The results remained very similar and only those from the GLM have been reported.

Following the survey a total of 23 nurses were recruited and participated in four focus groups (Operating theatre group = 8, Critical Care group = 9, ward A group = 3 and Ward B group = 3). A convenience sample method was used and participants were approached during a teaching session where information was provided and the purpose of the focus group was explained. The focus group discussions were arranged during a lunch time. All participants consented to participate in the focus groups. These participants had not participated in the survey and therefore had not completed the SCQ. Open questions were used and the main topics discussed were, their overall understanding of safety climate or culture, communication and manager commitment to safety, as these dimensions of safety climate had been found to be different between groups in the first phase of the study. Each focus group was facilitated by one researcher who acted as facilitator, and an observer who noted group dynamics and timed the session. The groups lasted between 40 to 50 minutes and were recorded and later transcribed. A six phase approach to a thematic analysis was undertaken. [34] The transcribed discussions were imported into NVIVO 10 for windows to facilitate the development of codes. In-vivo coding was utilised for first order coding, as using the participants own words provided a much closer interpretation of their voice in the coding process. [35] The initial codes were refined throughout the process of analysis and codes were checked back to the transcripts to ensure that the meaning of the code was valid in the context of the transcript. During second order coding the initial codes were reviewed and grouped into categories and eventually into sub-themes and themes. A process of checking coding between researchers was undertaken to ensure reliability and validity of the coding process.

RESULTS

Survey results

A total of 563 questionnaires were distributed and 319 questionnaires were completed and returned (response rate = 57%). Four questionnaires were excluded from the final analysis because they were completed by nurses who did not fulfil the selection criteria, i.e., not a band 5 or 6 adult nurse working in an acute hospital setting. Little's MCAR indicated that the

missing data were missing completely at random and were unlikely therefore to unduly affect the results (Little's MCAR test: $\chi^2 = 2368.11$, df = 2292, p = 0.131)

Table one illustrates the demographic data of the participants according to the specialist areas they worked in. There were more participants from critical care units than from other groups. The group identified as other included participants who stated that they worked in acute hospital setting areas such as, out patients, care of the elderly, oncology and haematology. The numbers of participants in these areas was low so these were grouped together.

Table One: Mean and standard deviation (SD) χ^2 and for demographic data for critical care, operating theatres, medicine, surgery and other clinical areas.

	Critical Care (n = 107)	Operating theatres (n = 49)	Medicine (n = 70)	Surgery (n= 54)	Other (n=24)	χ^2
Present position Mean (SD) years	3.12 (2.60)	4.26 (3.58)	3.69 (3.35)	3.12 (2.22)	3.21 (2.48)	p = 0.442
Years qualified Mean (SD) years	7.63 (5.47)	8.90 (6.85)	8.14 (6.46)	6.93 (6.00)	8.60 (6.04)	p = 0.317
Years Specialism Mean (SD) years	4.30 (3.77)	6.34 (5.25)	5.02 (3.84)	4.29 (3.74)	4.13 (2.70)	p = 0.195
Specialist qualification Percentage	50% (54/107)	43% (20/49)	37% (26/70)	33% (18/54)	58% (14/24)	P = 0.029*
Safety training Percentage	71% (76/107)	55% (27/49)	69% (48/70)	59% (32/54)	67% (16/24)	P = 0.032*

* statistically significant difference

Across the groups the participants had been working in their present position between 3 to 4 years. There was more variability across the groups in terms of how long the participants had been qualified with the critical care and surgery ward nurses being qualified as a registered nurse for less time. There was some variation in the amount of time the participants had been working within the specialism and the results indicate that the participants had been working in other areas before finally working within their specialist areas. The percentage of those reporting having undergone safety training ($\chi^2 = 6.12$, df = 4, p = 0.032) and those participants reporting having a specialist qualification ($\chi^2 = 9.83$, df = 4, p = 0.029) varied significantly across clinical specialism. All other variables did not vary significantly across

clinical specialism. All participants who had reported undergoing safety training undertaken in UK hospitals on an annual basis described this as mandatory training. Typically this includes training in manual handling, resuscitation and infection control.

Confirmatory Factor Analysis

The CFA goodness of fit measures indicated that there was a good model fit on some criteria with a significant Chi Square test (χ^2 = 1687.560, df = 824, p = < 0.001). Both the χ^2/df ratio of 2.05 and RMSEA value of 0.058 (90% CI interval 0.054 to 0.062) indicated a good model fit. However, the CFI was 0.805, although this was towards the higher end of the CFI range (0 to 1) it was below the acceptable threshold level (CFI >0.9) and suggests that the model could be improved.

The CFA regression weights (factor loadings) were similar to those from the original petroleum industry study (see supplementary table). However, there were four items that were particularly low and related to the dimensions of supportive environment, personal appreciation of risk and work environment. In relation to a supportive environment the item relating to, 'A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate', had a regression weight of 0.150 and the item relating to, 'When people ignore safety procedures here I feel it is none of my business', had a regression weight of 0.291. In the dimension of personal appreciation of risk, the item, 'I am rarely worried about being injured in the job', had a regression weight of 0.110 and in the dimension of Work environment the item, 'This is a safer place to work than other Trusts I have worked for', had a regression weight of 0.270. These items may not make a significant contribution to the perception of safety climate in a nursing population. Cox and Cheyne [7] kept lower regression weighted items in their original questionnaire and suggested that these items should be used with caution.

Cronbach's alpha was greater than 0.70 for five of the nine dimensions (Management commitment 0.84, Priority of safety 0.76, Communication 0.70, Personal priorities and need for safety 0.72, Work environment 0.72). There were four dimensions with a Cronbach's alpha of less than 0.70 (Safety rules 0.67, Supportive environment 0.55, Involvement 0.58, Personal appreciation of risk 0.48). There was some marginal improvement in Cronbach's

alpha when items with standardized regression weights of less than 0.3 were excluded (Supportive environment 0.55 to 0.57, Personal appreciation of risk 0.48 to 0.50, Work environment 0.72 to 0.74).

Comparison of safety climate scores

Following the CFA the factor scores derived from the survey were used to go onto explore differences in safety climate scores between nurses working in different clinical specialisms. Comparisons were made between nurses working in critical care areas, operating theatres, medical wards, surgical wards and other acute hospital settings as described above. Table two shows the adjusted GLM mean, 95% confidence interval by clinical specialism and F statistic, for all of the safety climate dimensions. Overall the scores were towards the higher range on the safety climate scale and suggested that participants reported a fairly positive safety climate for most of the dimensions. However, the work environment factor had lower scores across all the groups whilst personal priority of safety scored highly across all groups. There was a statistically significant difference between groups for Management Commitment ($F [4,266] = 4.66$, $p = 0.001$) and for Communication ($F [4,266] = 2.62$, $p = 0.035$).

Table Two: Comparison of the nine safety climate dimensions across clinical specialism adjusting for profile variable.

m = significantly different from Medicine; s= significantly different from surgery; o= significantly different from operating theatres

		Critical Care	Operating theatres	Medical wards	Surgical wards	Other	F test, p
Management Commitment	Mean	3.48	3.27[m,s]	3.75[o]	3.66[o]	3.31	F (4,266) = 4.66, p= 0.001
	(95% CI)	(3.34, 3.62)	(3.07, 3.67)	(3.59, 3.91)	(3.47, 3.85)	(2.99, 3.63)	
Priority of safety	Mean	3.54	3.44	3.73	3.50	3.61	F (4,266) = 1.29, p = 0.27
	(95% CI)	(3.39, 3.69)	(3.22, 3.66)	(3.55, 3.91)	(3.30, 3.71)	(3.26, 3.96)	
Communication	Mean	3.19	3.17	3.50	3.35	3.13	F (4,266) = 2.62, p= 0.035
	(95% CI)	(3.04, 3.33)	(2.96, 3.38)	(3.33, 3.67)	(3.15, 3.54)	(2.79, 3.47)	
Safety rules	Mean	3.18	3.23	3.43	3.40	2.90	F (4.266) = 1.96, p = 0.10
	(95% CI)	(3.01, 3.36)	(2.98, 3.48)	(3.22, 3.64)	(3.17, 3.64)	(2.49, 3.31)	
Supportive environment	Mean	3.66	3.67	3.86	3.75	3.63	F (4,266) = 1.85, p = 0.12
	(95% CI)	(3.55, 3.76)	(3.51, 3.82)	(3.73, 3.98)	(3.60, 3.89)	(3.38, 3.88)	
Involvement in safety	Mean	3.31	3.45	3.50	3.63	3.37	F (4,266) = 1.87, p = 0.12
	(95% CI)	(3.16, 3.46)	(3.24, 3.66)	(3.33, 3.68)	(3.43, 3.82)	(3.03, 3.71)	
Personal priorities and need for safety	Mean	4.20	4.31	4.37	4.33	4.11	F(4,266) = 1.89, p = 0.11
	(95% CI)	(4.10, 4.30)	(4.16, 4.45)	(4.25, 4.48)	(4.20, 447)	(3.88, 4.34)	
Personal appreciation of risk	Mean	3.19	3.15	3.36	3.44	3.35	F (4,226) = 0.92, p = .080
	(95% CI)	(3.05, 3.32)	(2.96, 3.34)	(3.20, 3.52)	(3.26, 3.61)	(3.04, 3.65)	
Work environment	Mean	2.62	2.65	2.68	2.82	2.85	F (4,266) = .092, p = 0.45
	(95% CI)	(2.47, 2.77)	(2.44, 2.86)	(2.50, 2.85)	(2.62, 3.02)	(2.51, 3.20)	

A Bonferroni post-hoc test revealed that there was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses (Mean = 3.27, 95% CI 3.07 – 3.47), compared with nurses working in medical wards (Mean = 3.75, 95% CI 3.59 – 3.91) and surgical ward settings (Mean = 3.66, 95% CI 3.47 – 3.85). Although there was significant variation in safety climate scores for communication across clinical specialism, none of the pair-wise comparisons achieved statistical significance at the 5% level, although the difference between critical care (Mean = 3.19, 95% CI 3.04 – 3.33) and the medical wards (Mean = 3.50 95% CI 3.33 – 3.67) came close (p = 0.056).

Thematic analysis

The results of the cross-sectional survey indicated a difference between nurses on the dimensions of Management commitment and though not statistically significant, Communication. During the focus groups participants were invited to discuss their understanding of safety culture and for their views of management and communication related to safety. Specific details of the differences found in the survey were not disclosed to the participants in order not to lead the discussion. Though these two aspects were discussed several other issues were also raised by participants. Three main themes emerged from the thematic analysis of the focus group data. These were Human Factors, Clinical Management and Protecting Patients. A further meta-theme was also identified as The System and Human Side of Caring.

Human Factors

The theme of Human Factors related to aspects of the environment such as design and staffing, the use of checklists and incident reporting. Aspects of physical environment were viewed as carrying potential risks and hazards to patients and the nurse is important in constantly checking equipment to ensure safety. For example, this participant stated that, *'I have to go round everywhere, checking the emergency crash call, check the monitors. The date they were serviced.'* (Critical Care group). Other participants recognised environmental design that has improved patient safety, such as, laminated flooring, '*We have a laminated grip flooring. They can still have a fall but it is much better for them.'* (Medical ward group).

The ratio of the numbers of patients to nurses was a concern, for example, *'Even in the current era, the ratio of nurses to patients is still a bit high. In terms of care, sometimes we are under so much pressure.'* (Medical Surgical group). All the groups mentioned the use of checklists. The operating theatre group mentioned the use of the World Health Organisation (WHO) checklist and the ward groups mentioned the use of intentional rounding. Though the content of these checklists are different they were seen as having advantages for patient safety and have been embedded in nursing practice. For example, '*We're very serious about protocols and policies as well and….we live by the checklist now*.' (Operating theatre group) and '*We have a checklist now and we check every single patient on the ward is safe.'* (Medical Surgical ward group).

There were ambivalent feelings regarding the use of incident reports where some participants viewed them as positive opportunities to learn from error, for example, '*You can learn from error, you can see it's not about blame culture.*' (Medical ward group), or were seen negatively, as this participant articulated, *'Yeah, a weapon, not something to help you. We're going to tell on you.*'(Critical care group). All of these approaches are systematic ways of managing error that are evident in nursing practice and the participants recognised the importance of these approaches to patient safety.

Clinical Management

The theme of Clinical Management related to communication processes and management behaviours that were relevant to the day-to-day management of patient care. Structured approaches to communication, such as handover, team briefing, and ward rounds were viewed as important for patient safety. Generally communication between nursing teams was seen as positive but communication between professions was identified as problematic, *'I think communication between nurses is good and between doctors and doctors is very good, but I think that there is a massive communication breakdown in people from different professions….I think information is lost all the time*.'(Critical care group) The role of the medical notes was viewed as being very important in communicating medical decisions to nursing staff but this was problematic for many participants. For example, *'Sometimes you are on night shift and you handover to the nurse who is taking over in the morning and you handover things that have happened and there's nothing written in the notes, nothing written*

by the doctors.' (Critical care group). The nurses perceived medical staff as not understanding the significance of the medical record for safe nursing care.

Manager behaviour was also identified as very important for the participants feeling supported in patient safety. Managers who were seen as approachable and proactive in managing patient safety were generally viewed as providing support for example, '*My manager tends to pay a lot of attention to those small details where the chart is not updated, he will remind staff, so he is very picky on the small things, which is good because it reminds everybody about what you are doing.*' (Medical Surgical group). Those managers who were seen as unsupportive tended to be reactive and not supportive of staff, for example, '*Just telling me what to do. It's just like another surgeon telling me what to do.'* (Operating theatre group).

Protecting patients

Protecting patients was a key theme that emerged as being important aspect of nursing practice relating to patient safety. This focused upon how nursing skill is applied to patient care and acting as a gatekeeper and advocate for patients. There was an overall sense that patients are vulnerable, for example, '*The nature of our patients we're receiving acutely unwell patients who are suffering from delirium and are vulnerable.*' (Medical ward group). There was a sense that nurses protect patients by ensuring safety whilst undertaking nursing tasks, for example, *'Administrating medication is a major thing and I think safety should be ensured all the time and I see we always check, because you've got a critically ill patient and the last thing you want is a drug error.'* (Critical care group). There was also a sense that nurses need to challenge others. For example, *'I think when it comes to patient safety everyone has to take responsibility for safety, the doctors just don't do it. We encourage, we try to make everyone to be attentive but you have to challenge them.'* (Critical care group). There was a clear sense that the participants felt that they had a role in protecting patients from harm.

Joint synthesis of survey and focus group findings

The results of the cross-sectional survey found a variation in the dimension of Communication between nurses working in critical care and medical wards, though pairwise

comparisons were not statistically significant. Table three shows the mean and 95% confidence intervals for the dimension of Communication and a summary of the themes identified in the thematic analysis of the focus group discussions. The ward focus groups identified nurse-to-nurse communication as important for patient safety and these groups had slightly higher safety climate scores in this area. The Critical Care and Operating theatre focus group highlighted challenges associated with nurse to doctor communication.

Table Three: Differences in the dimension of communication between critical care, operating theatres, medical and surgical wards for the SCQ and theme

SCQ Communication score	Summary of thematic analysis
Critical care Mean = 3.19 (3.04, 3,33)	The main mechanism for communication was the ward round. Problems were identified where communication was poor following a ward round or where medical staff did not record in the medical record.
Operating theatres Mean = 3.17 (2.96, 3.38)	The main mechanism for communication was the WHO checklist and team briefing. There were challenges associated with compliance with these approaches from surgeons.
Medical wards Mean = 3.50 (3.33, 3.67) Surgical wards Mean = 3.35 (3.15, 3.54)	The main focus of communication was related to handovers between nursing teams and ward rounds. These seem to work well.

There was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses, compared with nurses working in medical and surgical ward settings, with operating theatres having a lower score for Management Commitment. Table four shows the mean and confidence intervals for the dimension of Management Commitment and the themes that were identified in the focus groups. The operating theatre group reported more reactive and unsupportive manager behaviours in the focus group discussion. Whereas, the other areas generally reported proactive and supportive

manager behaviours in the focus groups, the operating theatre focus group reported reactive style of management.

Table Four: Differences in the dimension of management commitment to patient safety, between critical care, operating theatres, medical and surgical wards for the SCQ and themes.

SCQ Management commitment score	Summary of thematic analysis
Critical care Mean = 3.48 (3.34, 3.62)	Being approachable and accessible to support staff. Having experience and clinical credibility.
Operating theatres Mean = 3.27 (3.07, 3.67)	The perception that manager take sides with medical staff, not providing help and advice to nurses when they approach managers for assistance, and having an agenda related to targets, managers side with the surgeons and do not support the nursing staff, that the rules do not apply to surgeons.
Medical wards Mean = 3.75 (3.39, 3.91) Surgical wards Mean = 3.66 (3.47, 3.85)	Being proactive in supporting patient safety and reminding staff about compliance to safety procedures. Working clinically in the area and having clinical credibility with the nursing staff was highly valued and being approachable and accessible to nursing staff when they feel that they need support with problems related to patient safety.

The System and Human Side of Caring

A meta-theme, or overarching theme was identified from the three main themes and was labelled, the system and the human side of caring. This holistic view of the data captures two aspects of patient safety that seemed to be apparent within the data. That is, the system in which caring takes place, and this includes the physical environment, the design of that environment, and the system processes that have been put in place to assist patient safety with the use of checklists and incident reporting. These systematic organisational structures and processes provide the backdrop and the context in which caring takes place. The human side of caring includes the personal and the interpersonal aspects of care, the need to communicate within nursing teams and to handover care to each other. The relationship with clinical

managers was important to provide support for safe clinical care. The importance of interaction with other disciplines and the problems associated with that was a key component. Finally, the acknowledgement of patient's vulnerability within the system, and that nurses feel it is an important aspect of their role to act as an advocate and to protect patients through acting as a gatekeeper. Safety lies within an interaction between these two aspects of the clinical environment.

DISCUSSION

The application of High Reliability Organisation theory has underpinned the approach to patient safety in the past decade in the UK.[12] and the introduction of Human Factor approaches to patient safety is high on the agenda in the UK at the present time. The results of this study indicates that though human factor approaches are an important aspect of safe nursing practice, these approaches need to be supported with communication and management behaviours that rely upon good interpersonal skills. The emergent meta-theme of the system and the human side of caring indicates that attitudes and organisational culture are shaped and developed within the context of the transpersonal and the results indicate that support and communication empower nurses to advocate and protect their patients. The advent and development of checklists, the implementation of human factor and high reliability approaches are important and these have had a significant impact on patient safety but this study highlights other aspects of social behaviour and communication that can have an impact on patient safety. Indeed, too much focus upon targets and processes can be counterproductive. [36]

The SCQ has been used in the NHS extensively, however, the factor structure had not validated within a healthcare population before its use. The results of the confirmatory factor analysis undertaken here with a nursing sample indicated that the SCQ did have an acceptable level of model fit for some but not all criteria. The main focus of this study was to explore and understand variation in safety climate between specialisms and the SCQ provided some measurement that enabled further exploration of this variation. However, further work needs to be undertaken to fully validate this tool in the healthcare context. This tool was used extensively in the NHS without confirmation of its factor structure and these results illustrate

that it is important to ensure that tools developed in one context are evaluated for fit into another context.

The findings indicated there was a lower safety climate in operating theatres compared to ward areas for management commitment. Both critical care and operating theatre groups also scored lower for communication than medical ward areas, though this was close to, but not statistically different. This may seem surprising, given that in recent years there has been widespread introduction of High Reliability Organisation approaches into critical care units and operating theatres, such as the WHO checklist into operating theatres and the introduction of reliability and standardisation measures in intensive care units. [37-38]. However, these results are consistent with results from other countries and may indicate that there is a fundamental difference in safety climate in different clinical settings and it has been suggested that these differences are associated with the severity or complexity of the patient condition, high patient turnover, or the technological complexity of the care delivered.[16-18] The results of this mixed methods study may point to other factors associated with management and communication differences in these areas rather than the highly technical aspects of patient care associated with critical areas. It is interesting to note that the SCQ does not stipulate whether management commitment indicates middle or senior management. It was clear in the focus group discussions that nurses see their ward or unit manager as their manager. How nurses interpret these issues has implications for how safety climate scores can be interpreted.

In a post Francis Inquiry [36] era, nursing care in particular has had increasing scrutiny of its practice, and these results indicate that there is a focus on safety in clinical practice and this is reflected in the perceptions and attitudes of the nurses who participated in this study. The factor scores of Personal priorities and need for safety, were consistently high across all groups, suggesting that for the participants, safety is an important priority in patient care for these nurses and this was reflected in the focus group discussions. The factor scores for Work environment were consistently low across all groups and the focus group discussions highlighted the availability of equipment, staffing, the resources and time available to undertake the work are important aspects of safety in nursing practice.

It is acknowledged that the results need to be considered in the light of a cross-sectional survey response rate of 57% and the fact that the number of participants in some of the focus groups was low. However, the response rate is similar to other work undertaken in the field and although there were low numbers in some focus groups robust data was generated. The sample in this study comprised of registered nurses undertaking post-qualification education and these nurses are highly motivated to undertake professional development. The utilisation of a mixed methods design does not indicate direct association between the quantitative and qualitative aspects of the study. However, the results of this study raise some important issues relating to the underlying drivers of safety climate in nursing practice and the importance of using a mixed methodology to provide a deeper insight into the mechanisms driving safety climate in nursing.

ACKNOWLEDGEMENTS

To the RCN for granting permission to use their version of the SCQ. To the participants who gave their time to complete the SCQ.

REFERENCES

1. The Health Foundation. Measuring Safety Culture. London: The Health Foundation 2011.
2. The Health Foundation. The measurement and monitoring of safety. London: The Health Foundation. 2013.
3. Health and Safety Commission (HSC). Organising for Safety: Third Report. ACSNI Study Group on Human Factors. Sudbury: HSE Books. 1993 P15.
4. Health and Safety Executive. Reducing error and influencing behaviour. London: HSE. 1999.
5. Pronovost P and Sexton B. Assessing safety culture: guidelines and recommendations. *Qual Saf Health Care*. 2005;14:231–233.
6. Freeth D, Sandall J, Allan T, et al. A methodological study to compare survey-based and observation-based evaluations of organisational and safety cultures and then compare both approaches with makers of the quality of care. Health Technology Assessment. 2012;6:25.
7. Cox SJ and Cheyne AJT. Assessing safety culture in offshore environments. *Saf Sci.* 2000;34:111–129.
8. Currie L and Watterson L. Measuring the safety climate in NHS organisations. *Nurs Stand*. 2010;24(24):35–38.
9. University of Loughborough. Safety climate measurement user guide and toolkit. 2000. Available at: http://www.lboro.ac.uk/departments/sbe/downloads/pmdc/safety-climate-assessment-toolkit.pdf (assessed July 2013)
10. The Health Foundation. High Reliability Organisations. London: The Health Foundation. 2011.
11. Health and Safety Executive. High reliability organisations: A review of the literature. London: HSE. 2011.
12. Department of Health. An Organisation with a memory. London: The Stationary Office. 2001.
13. Olsen E and Aase KA. Comparative study of safety climate differences in healthcare and the petroleum industry. *Qual Saf Health Care.* 2010;19(Suppl 3):175–179.
14. Currie L and Cooper C. Measuring staffs' perceptions of safety climate: testing The Safety Climate Questionnaire (SCQ) for use in healthcare. London: RCN. 2008.
15. Zohar D, Livne Y, Tenne-Gazit O, et al. Healthcare climate: a framework for measuring and improving patient safety. *Critical Care Medicine.* 2007;35:1312–1317.

16. Hartmann CW, Rosen AK, Meterfo M, et al. An overview of patient safety climate in the VA. *Health Services Research.* 2008;43(4):1263–1284.
17. Singer SJ, Gaba DM, Fulwell A, et al. Patient safety climate in 92 US hospital: differences by work area and discipline. *Medical Care.* 2009;47:23–31.
18. Campbell EG, Singer S, Kitch BT, et al. Patient Safety Climate in Hospitals: Act Locally on Variation Across Units. *Joint Commission on Accreditation of Healthcare Organizations.* 2010;36(7):316–326.
19. Abbas HAE, Bassiuni NA, Baddar FM Perception of front-line healthcare providers toward patient safety: A preliminary study in a university hospital in Egypt. Topics in Advanced Practice Nursing e-journal. 2008:8(2): Medscape.
20. Abdou HA and Saber KM A baseline assessment of patient safety culture among nurses at student university hospital. World Journal of Medical Sciences. 2011:6(1): 17 – 26.
21. Needleman J, Buerhaus P, Pankratz VS, et al. Nurse staffing and inpatient hospital mortality. *N Engl J Med.* 2011;364:1037-1045.
22. Aiken LH, Sermeus W, Van den Heede K et al Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. British Medical Journal. 2012:344 e1717.
23. Agnew C, Flin RH, Mearns KJ. (2013). Patient safety climate and worker safety behaviours in acute hospitals in Scotland. Journal of Safety Research. 2013: 45: 95-101.
24. Leech N and Onwuegbuzie AJ (2009) A typology of mixed methods research designs. Quality and Quantity. 43: 265 – 275
25. Glynos J and Howarth D (2007) Logics of Critical Explanation in Social and Political Theory. London: Routledge.
26. Guetterman TC, Fetters MD, Creswell JW Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med. 2015 Nov; 13(6): 554–561.
27. Comrey, AL and Lee, HB*A. First Course in Factor Analysis.* Hillsdale, NJ: Erlbaum. 1992.
28. Elliot AC, Hynan LS, Reisch S, et al. Preparing data for analysis using Microsoft Excel. *J Investig Med.* 2006;54:6:334–341.
29. Little RJA and Rubin DB. Statistical analysis with missing data (2nd ed.). New York: Wiley. 2002.
30. Kline P. An easy guide to factor analysis. London: Routledge. 2008.

31. Hu L and Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*.1999;6(1):1-55.
32. Steiger JH. Understanding the limitations of global fit assessment in structural equation modelling. *Individual Differences.* 2007;42:893–898.
33. Lui RY. Bootstrap procedures under some Non-I.I.D. models. The Annals of Statistics. 1988;4:1696-1708.
34. Braun V and Clarke V Using thematic analysis in psychology. Qualitative Research in Psychology. 2006.3 (2): 77-101.
35. Saldaňa J The coding manual for qualitative researchers. (2nd Ed). 2013. London: Sage Publications.
36. Francis R Report of the Mid-Staffordshire NHS Foundation Trust public inquiry. 2013 London: The Stationary Office.
37. Haynes AB, Weiser TG, Berry WR, et al. (2009) A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;360:491-499.
38. Pronovost PJ, Berenholtz SM, Goeschel CA. Creating High Reliability in Health Care Organizations. *Health Services Research.* 2006;41(4) Part II:1599–1617.

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Management Commitment			
Managers and supervisors express concern if safety procedures are not followed	0.440	0.633	0.696
In my workplace managers/supervisors show interest in my safety	0.520	0.689	0.727
In my workplace management turn a blind eye to safety issues	0.737	0.618	0.712
In my workplace management acts quickly to correct safety problems	0.811	0.743	0.806
Corrective action is always taken when management is told about unsafe practice	0.690	0.527	0.571
Management acts only after incidents have occurred	0.500	0.666	0.710
Management acts decisively when a safety concern is raised	0.792	0.668	0.732
Priority of Safety			
Management considers safety to be equally important as getting the work done	0.534	0.717	0.757
Safety rules and procedures are carefully followed	0.585	0.592	0.670
Management clearly considers the safety of staff of great importance	0.665	0.659	0.688
I believe that safety issues are not assigned a high priority	0.585	0.675	0.742
Communication			
There is good communication here about safety issues which affect me	0.731	0.596	0.653
Safety information is always brought to my attention by my line manager/supervisor	0.633	0.596	0.638
I do not receive praise for working safely	0.481	0.511	0.569
My line manager/supervisor does not always inform me of current concern and issues	0.594	0.636	0.703
Management operates an open door policy on safety issues	0.541	0.511	0.574
Safety rules			
Some safety rules and procedures do not need to be followed to get the job done safely	0.724	0.600	0.580
Some health and safety rules and procedures are not really practical	0.685	0.622	0.744
Sometimes it is necessary to depart from safety requirements in order to get the work done	0.583	0.699	0.709
Supportive environment			
I can influence health and safety here	0.543	0.460	0.479
A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate	0.367	0.150*	0.153*
I am strongly encouraged to report unsafe conditions	0.639	0.692	0.756
When people ignore safety procedures here I feel it is none of my business	0.480	0.291*	0.387*
Co-workers often give tips to each other on how to work safely	0.323	0.452	0.498
Employees are not encouraged to raise safety concerns	0.421	0.508	0.624

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Personal priorities and need for safety			
A safe place to work has a lot of personal meaning to me	0.571	0.519	0.600
It is important to me that there is a continuing emphasis on safety	0.655	0.560	0.625
I understand the safety rules for my job	0.642	0.664	0.832
Personally I feel that safety issues are not the most important aspect of my job	0.500	0.571	0.694
Safety is the number one priority in my mind when completing a job	0.623	0.617	0.685
Personal appreciation of risk			
I am clear about what my responsibilities are for health and safety	0.273	0.561	0.744
I am sure it is only a matter of time before I am involved in an incident	0.782	0.382	0.383
In my workplace the chance of being involved in an incident are high	0.464	0.548	0.570
I am rarely worried about being injured on the job	0.286	0.110*	0.086*
Involvement			
I am involved with safety issues at work	0.687	0.657	0.602
I am never involved in the ongoing review of safety	0.524	0.402	0.541
I am involved in informing management of important safety issues	0.724	0.671	0.677
Work environment			
There are always enough people available to get the job done safely	0.596	0.527	0.593
Sometimes I am not given enough time to get the job done safely	0.668	0.610	0.543
Sometimes conditions here hinder my ability to work safely	0.666	0.724	0.727
Operational targets often conflict with safety measures	0.795	0.590	0.662
I cannot always get the equipment I need to do the job safely	0.448	0.587	0.645
This is a safer place to work than other Trusts I have worked for	0.256	0.270*	0.392*

McGill

Mixed Methods Appraisal Tool (MMAT) – Version 2011

For dissemination, application, and feedback: Please contact pierre.pluye@mcgill.ca, Department of Family Medicine, McGill University, Canada.

The MMAT is comprised of two parts (see below): criteria (Part I) and tutorial (Part II). While the content validity and the reliability of the pilot version of the MMAT have been examined, this critical appraisal tool is still in development. Thus, the MMAT must be used with caution, and users' feedback is appreciated. Cite the present version as follows.

> Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'Cathain, A., Griffiths, F., Boardman, F., Gagnon, M.P., & Rousseau, M.C. (2011). *Proposal: A mixed methods appraisal tool for systematic mixed studies reviews*. Retrieved on [date] from http://mixedmethodsappraisaltoolpublic.pbworks.com. Archived by WebCite® at http://www.webcitation.org/5tTRTc9yJ

Purpose: The MMAT has been designed for the appraisal stage of complex systematic literature reviews that include qualitative, quantitative and mixed methods studies (mixed studies reviews). The MMAT permits to concomitantly appraise and describe the methodological quality for three methodological domains: mixed, qualitative and quantitative (subdivided into three sub-domains: randomized controlled, non-randomized, and descriptive). Therefore, using the MMAT requires experience or training in these domains. E.g., MMAT users may be helped by a colleague with specific expertise when needed. The MMAT allows the appraisal of most common types of study methodology and design. For appraising a qualitative study, use section 1 of the MMAT. For a quantitative study, use section 2 or 3 or 4, for randomized controlled, non-randomized, and descriptive studies, respectively. For a mixed methods study, use section 1 for appraising the qualitative component, the appropriate section for the quantitative component (2 or 3 or 4), and section 5 for the mixed methods component. For each relevant study selected for a systematic mixed studies review, the methodological quality can then be described using the corresponding criteria. This may lead to exclude studies with lowest quality from the synthesis, or to consider the quality of studies for contrasting their results (e.g., low quality vs. high).

Scoring metrics: For each retained study, an overall quality score may be not informative (in comparison to a descriptive summary using MMAT criteria), but might be calculated using the MMAT. Since there are only a few criteria for each domain, the score can be presented using descriptors such as *, **, ***, and ****. For qualitative and quantitative studies, this score can be the number of criteria met divided by four (scores varying from 25% (*) -one criterion met- to 100% (****) -all criteria met-). For mixed methods research studies, the premise is that the overall quality of a combination cannot exceed the quality of its weakest component. Thus, the overall quality score is the lowest score of the study components. The score is 25% (*) when *QUAL=1 or QUAN=1 or MM=0*; it is 50% (**) when *QUAL=2 or QUAN=2 or MM=1*; it is 75% (***) when *QUAL=3 or QUAN=3 or MM=2*; and it is 100% (****) when *QUAL=4 and QUAN=4 and MM=3* (QUAL being the score of the qualitative component; QUAN the score of the quantitative component; and MM the score of the mixed methods component).

Rationale: There are general criteria for planning, designing and reporting mixed methods research (Creswell and Plano Clark, 2010), but there is no consensus on key specific criteria for appraising the methodological quality of mixed methods studies (O'Cathain, Murphy and Nicholl, 2008). Based on a critical examination of 17 health-related systematic mixed studies reviews, an initial 15-criteria version of MMAT was proposed (Pluye, Gagnon, Griffiths and Johnson-Lafleur, 2009). This was pilot tested in 2009. Two raters assessed 29 studies using the pilot MMAT criteria and tutorial (Pace, Pluye, Bartlett, Macaulay et al., 2010). Based on this pilot exercise, it is anticipated that applying MMAT may take on average 15 minutes per study (hence efficient), and that the Intra-Class Correlation might be around 0.8 (hence reliable). The present 2011 revision is based on feedback from four workshops, and a comprehensive framework for assessing the quality of mixed methods research (O'Cathain, 2010).

Conclusion: The MMAT has been designed to appraise the *methodological quality* of the studies retained for a systematic mixed studies review, not the quality of their *reporting* (writing). This distinction is important, as good research may not be 'well' reported. If reviewers want to genuinely assess the former, companion papers and research reports should be collected when some criteria are not met, and authors of the corresponding publications should be contacted for additional information. Collecting additional data is usually necessary to appraise *qualitative research and mixed methods studies*, as there are no uniform standards for reporting study characteristics in these domains (www.equator-network.org), in contrast, e.g., to the CONSORT statement for reporting randomized controlled trials (www.consort-statement.org).

Authors and contributors: Pierre Pluye[1], Marie-Pierre Gagnon[2], Frances Griffiths[3] and Janique Johnson-Lafleur[1] proposed an initial version of MMAT criteria (Pluye et al., 2009). Romina Pace[1] and Pierre Pluye[1] led the pilot test. Gillian Bartlett[1], Belinda Nicolau[4], Robbyn Seller[1], Justin Jagosh[1], Jon Salsberg[1] and Ann Macaulay[1] contributed to the pilot work (Pace et al., 2010). Pierre Pluye[1], Émilie Robert[5], Margaret Cargo[6], Alicia O'Cathain[7], Frances Griffiths[3], Felicity Boardman[3], Marie-Pierre Gagnon[2], Gillian Bartlett[1], and Marie-Claude Rousseau[8] contributed to the present 2011 version.

Affiliations: 1. Department of Family Medicine, McGill University, Canada; 2. Faculté des sciences infirmières, Université Laval, Canada; 3. Warwick Medical School, University of Warwick, UK; 4. Faculty of Dentistry, McGill University, Canada; 5. Centre de recherche du CHUM, Université de Montréal, Canada; 6. School of Health Sciences, University of South Australia, Australia; 7. Medical Care Research Unit, ScHARR, University of Sheffield, UK; 8. INRS-Institut Armand Frappier, Laval, Canada.

PART I. MMAT criteria & one-page template (to be included in appraisal forms)

Types of mixed methods study components or primary studies	Methodological quality criteria (see tutorial for definitions and examples)	Responses			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	• Are there clear qualitative and quantitative research questions (or objectives*), or a clear mixed methods question (or objective*)?				Page 5
	• Do the collected data allow address the research question (objective)? E.g., consider whether the follow-up period is long enough for the outcome to occur (for longitudinal studies or study components).				Page 5
	Further appraisal may be not feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.				
1. Qualitative	1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?				Page 8
	1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?				Page 8
	1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?				Page 8 & Page 18
	1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?				Page 8
2. Quantitative randomized controlled (trials)	2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?				N/A
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)?				N/A
	2.3. Are there complete outcome data (80% or above)?				N/A
	2.4. Is there low withdrawal/drop-out (below 20%)?				N/A
3. Quantitative non-randomized	3.1. Are participants (organizations) recruited in a way that minimizes selection bias?				N/A
	3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?				N/A
	3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?				N/A
	3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?				N/A
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?				Page 6
	4.2. Is the sample representative of the population understudy?				Page 6
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?				Page 6
	4.4. Is there an acceptable response rate (60% or above)?				Page 8 & page 18
5. Mixed methods	5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?				Page 6
	5.2. Is the integration of qualitative and quantitative data (or results*) relevant to address the research question (objective)?				Page 14 - 16
	5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results*) in a triangulation design?				Page 18
	Criteria for the qualitative component (1.1 to 1.4), and appropriate criteria for the quantitative component (2.1 to 2.4, or 3.1 to 3.4, or 4.1 to 4.4), must be also applied.				

*These two items are not considered as double-barreled items since in mixed methods research, (1) there may be research questions (quantitative research) or research objectives (qualitative research), and (2) data may be integrated, and/or qualitative findings and quantitative results can be integrated.

PART II. MMAT tutorial

Types of mixed methods study components or primary studies	Methodological quality criteria
1. Qualitative Common types of qualitative research methodology include: A. Ethnography The aim of the study is to describe and interpret the shared cultural behaviour of a group of individuals. B. Phenomenology The study focuses on the subjective experiences and interpretations of a phenomenon encountered by individuals. C. Narrative The study analyzes life experiences of an individual or a group. D. Grounded theory Generation of theory from data in the process of conducting research (data collection occurs first). E. Case study In-depth exploration and/or explanation of issues intrinsic to a particular case. A case can be anything from a decision-making process, to a person, an organization, or a country. F. Qualitative description There is no specific methodology, but a qualitative data collection and analysis, e.g., in-depth interviews or focus groups, and hybrid thematic analysis (inductive and deductive). Key references: Creswell, 1998; Schwandt, 2001; Sandelowski, 2010.	**1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?** E.g., consider whether (a) the selection of the participants is clear, and appropriate to collect relevant and rich data; and (b) reasons why certain potential participants chose not to participate are explained. **1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?** E.g., consider whether (a) the method of data collection is clear (in depth interviews and/or group interviews, and/or observations and/or documentary sources); (b) the form of the data is clear (tape recording, video material, and/or field notes for instance); (c) changes are explained when methods are altered during the study; and (d) the qualitative data analysis addresses the question. **1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?*** E.g., consider whether the study context and how findings relate to the context or characteristics of the context are explained (how findings are influenced by or influence the context). "For example, a researcher wishing to observe care in an acute hospital around the clock may not be able to study more than one hospital. (…) Here, it is essential to take care to describe the context and particulars of the case [the hospital] and to flag up for the reader the similarities and differences between the case and other settings of the same type" (Mays & Pope, 1995). The notion of context may be conceived in different ways depending on the approach (methodology) tradition. **1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?*** E.g., consider whether (a) researchers critically explain how findings relate to their perspective, role, and interactions with participants (how the research process is influenced by or influences the researcher); (b) researcher's role is influential at all stages (formulation of a research question, data collection, data analysis and interpretation of findings); and (c) researchers explain their reaction to critical events that occurred during the study. The notion of reflexivity may be conceived in different ways depending on the approach (methodology) tradition. E.g., "at a minimum, researchers employing a generic approach [qualitative description] must explicitly identify their disciplinary affiliation, what brought them to the question, and the assumptions they make about the topic of interest" (Caelli, Ray & Mill, 2003, p. 5).

*See suggestion on the MMAT wiki homepage (under '2011 version'): Independent reviewers can establish a common understanding of these two items prior to beginning the critical appraisal.

Types of mixed methods study components or primary studies	Methodological quality criteria
2. Quantitative randomized controlled (trials) Randomized controlled clinical trial: A clinical study in which individual participants are allocated to intervention or control groups by randomization (intervention assigned by researchers). Key references: Higgins & Green, 2008; Porta, 2008; Oxford Center for Evidence based medicine, 2009.	**2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?** In a randomized controlled trial, the allocation of a participant (or a data collection unit, e.g., a school) into the intervention or control group is based solely on chance, and researchers describe how the randomization schedule is generated. "A simple statement such as 'we randomly allocated' or 'using a randomized design' is insufficient". *Simple randomization:* Allocation of participants to groups by chance by following a predetermined plan/sequence. "Usually it is achieved by referring to a published list of random numbers, or to a list of random assignments generated by a computer". *Sequence generation:* "The rule for allocating interventions to participants must be specified, based on some chance (random) process". Researchers provide sufficient detail to allow a readers' appraisal of whether it produces comparable groups. E.g., blocked randomization (to ensure particular allocation ratios to the intervention groups), or stratified randomization (randomization performed separately within strata), or minimization (to make small groups closely similar with respect to several characteristics).
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)? *The allocation concealment protects assignment sequence until allocation.* E.g., researchers and participants are unaware of the assignment sequence up to the point of allocation. E.g., group assignment is concealed in opaque envelops until allocation. *The blinding protects assignment sequence after allocation.* E.g., researchers and/or participants are unaware of the group a participant is allocated to during the course of the study.
	2.3. Are there complete outcome data (80% or above)? E.g., almost all the participants contributed to almost all measures.
	2.4. Is there low withdrawal/drop-out (below 20%)? E.g., almost all the participants completed the study.

Types of mixed methods study components or primary studies	Methodological quality criteria
3. Quantitative non-randomized Common types of design include (A) non-randomized controlled trials, and (B-C-D) observational analytic study or component where the intervention/exposure is defined/assessed, but not assigned by researchers. A. Non-randomized controlled trials The intervention is assigned by researchers, but there is no randomization, e.g., a pseudo-randomization. A non-random method of allocation is not reliable in producing alone similar groups. B. Cohort study Subsets of a defined population are assessed as exposed, not exposed, or exposed at different degrees to factors of interest. Participants are followed over time to determine if an outcome occurs (prospective longitudinal). C. Case-control study Cases, e.g., patients, associated with a certain outcome are selected, alongside a corresponding group of controls. Data is collected on whether cases and controls were exposed to the factor under study (retrospective). D. Cross-sectional analytic study At one particular time, the relationship between health-related characteristics (outcome) and other factors (intervention/exposure) is examined. E.g., the frequency of outcomes is compared in different population sub-groups according to the presence/absence (or level) of the intervention/exposure. Key references for observational analytic studies: Higgins & Green, 2008; Wells, Shea, O'Connell, Peterson, et al., 2009.	**3.1. Are participants (organizations) recruited in a way that minimizes selection bias?** At recruitment stage: For cohort studies, e.g., consider whether the exposed (or with intervention) and non-exposed (or without intervention) groups are recruited from the same population. For case-control studies, e.g., consider whether same inclusion and exclusion criteria were applied to cases and controls, and whether recruitment was done independently of the intervention or exposure status. For cross-sectional analytic studies, e.g., consider whether the sample is representative of the population. **3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?** At data collection stage: E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) the measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure. For non-randomized controlled trials, the intervention is assigned by researchers, and so consider whether there was absence/presence of a contamination. E.g., the control group may be indirectly exposed to the intervention through family or community relationships. **3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** At data analysis stage: For cohort, case-control and cross-sectional, e.g., consider whether (a) the most important factors are taken into account in the analysis; (b) a table lists key demographic information comparing both groups, and there are no obvious dissimilarities between groups that may account for any differences in outcomes, or dissimilarities are taken into account in the analysis. **3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?**

Types of mixed methods study components or primary studies	Methodological quality criteria
4. Quantitative descriptive studies Common types of design include single-group studies: A. Incidence or prevalence study without comparison group In a defined population at one particular time, what is happening in a population, e.g., frequencies of factors (importance of problems), is described (portrayed). B. Case series A collection of individuals with similar characteristics are used to describe an outcome. C. Case report An individual or a group with a unique/unusual outcome is described in details. Key references: Critical Appraisal Skills Programme, 2009; Draugalis, Coons & Plaza, 2008.	**4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?** E.g., consider whether (a) the source of sample is relevant to the population under study; (b) when appropriate, there is a standard procedure for sampling, and the sample size is justified (using power calculation for instance).
	4.2. Is the sample representative of the population understudy? E.g., consider whether (a) inclusion and exclusion criteria are explained; and (b) reasons why certain eligible individuals chose not to participate are explained.
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)? E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure.
	4.4. Is there an acceptable response rate (60% or above)? The response rate is not pertinent for case series and case report. E.g., there is no expectation that a case series would include all patients in a similar situation.

Types of mixed methods study components or primary studies	Methodological quality criteria
5. Mixed methods Common types of design include: A. Sequential explanatory design The quantitative component is followed by the qualitative. The purpose is to explain quantitative results using qualitative findings. E.g., the quantitative results guide the selection of qualitative data sources and data collection, and the qualitative findings contribute to the interpretation of quantitative results. B. Sequential exploratory design The qualitative component is followed by the quantitative. The purpose is to explore, develop and test an instrument (or taxonomy), or a conceptual framework (or theoretical model). E.g., the qualitative findings inform the quantitative data collection, and the quantitative results allow a generalization of the qualitative findings. C. Triangulation design The qualitative and quantitative components are concomitant. The purpose is to examine the same phenomenon by interpreting qualitative and quantitative results (bringing data analysis together at the interpretation stage), or by integrating qualitative and quantitative datasets (e.g., data on same cases), or by transforming data (e.g., quantization of qualitative data). D. Embedded design The qualitative and quantitative components are concomitant. The purpose is to support a qualitative study with a quantitative sub-study (measures), or to better understand a specific issue of a quantitative study using a qualitative sub-study, e.g., the efficacy or the implementation of an intervention based on the views of participants. Key references: Creswell & Plano Clark, 2007; O'Cathain, 2010.	**5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?** E.g., the rationale for integrating qualitative and quantitative methods to answer the research question is explained. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question (objective)?** E.g., there is evidence that data gathered by both research methods was brought together to form a complete picture, and answer the research question; authors explain when integration occurred (during the data collection-analysis or/and during the interpretation of qualitative and quantitative results); they explain how integration occurred and who participated in this integration. **5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results)?**

References

- Caelli, K., Ray, L., & Mill, J. (2003). 'Clear as Mud': Toward greater clarity in generic qualitative research. *International Journal of Qualitative Methods, 2(*2), 1-23.
- Creswell, J., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. London: Sage.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks: Sage.
- Critical Appraisal Skills Programme (2009). CASP appraisal tools. Retrieved on August 26, 2009 from: *www.phru.nhs.uk/pages/PHD/resources.htm*
- Draugalis, J.R., Coons, S.J., & Plaza, C.M. (2008). Best practices for survey research reports: a synopsis for authors and reviewers. *American Journal of Pharmaceutical Education, 72(1),* e11.
- Higgins, J.P.T. & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* - Version 5.0.1 [updated September 2008]. The Cochrane Collaboration. Retrieved on August 26, 2009 from www.cochrane-handbook.org
- Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *British Medical Journal, 311(6997),* 109-112.
- O'Cathain, A., Murphy, E. & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research and Policy, 13(2),* 92-98.
- O'Cathain, A. (2010). Assessing the quality of mixed methods research: Towards a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research (2nd edition)* (pp. 531-555). Thousand Oaks: Sage.
- Pace, R., Pluye, P., Bartlett, G., Macaulay, A., Salsberg, J., Jagosh, J., & Seller, R. (2010). *Reliability of a tool for concomitantly appraising the methodological quality of qualitative, quantitative and mixed methods research: a pilot study*. 38th Annual Meeting of the North American Primary Care Research Group (NAPCRG), Seattle, USA.
- Pluye, P., Gagnon, M.P., Griffiths, F. & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46(4),* 529-46.
- Oxford Center for Evidence Based Medicine (2009). Levels of evidence. Retrieved on July 7, 2009 from www.cebm.net/levels_of_evidence.asp
- Porta, M. (2008). *A Dictionary of Epidemiology*. New York: Oxford University Press.
- Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Research in Nursing and Health, 33*(1), 77-84.
- Schwandt, T. (2001). *Dictionary of qualitive inquiry*. Thousand Oaks: Sage.
- Wells, G.A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2009). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses*. The Cochrane Non-Randomized Studies Method Group. Retrieved on July 7, 2009 from www.ohri.ca/programs/clinical_epidemiology/oxford.htm

BMJ Open

Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016977.R2
Article Type:	Research
Date Submitted by the Author:	18-Aug-2017
Complete List of Authors:	Tarling, Maggie; City University School of Health Sciences, Nursing Jones, Anne; Kings College London, Adult Nursing Murrells, Trevor; Kings College London, National Nursing Research Unit McCutcheon, Helen; University of Queensland, Faculty of Health and Behavioural Sciences
Primary Subject Heading:	Nursing
Secondary Subject Heading:	Communication, Health services research, Intensive care, Qualitative research
Keywords:	Health & safety < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Risk management < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Adult intensive & critical care < INTENSIVE & CRITICAL CARE, QUALITATIVE RESEARCH

Title: Comparing safety climate for nurses working in operating theatres, critical care and ward areas in the UK: A mixed methods study.

Corresponding author:

Maggie Tarling
School of Health Sciences
City University London
Northampton Square
London, EC1 0HB
United Kingdom

Email: Maggie.Tarling.2@city.ac.uk
T: +44 (0)20 7040 0201

Co-authors:

Anne Jones
Florence Nightingale Faculty of Nursing & Midwifery
King's College London
57 Waterloo Road
London, SE1 8WA
United Kingdom

Trevor Murrells
Florence Nightingale Faculty of Nursing & Midwifery
King's College London
57 Waterloo Road
London, SE1 8WA
United Kingdom

Helen McCutcheon
The University of Queensland
Brisbane St Lucia, QLD 4072
Australia

Keywords: patient safety; safety climate; nursing; mixed methods; human factors

Word count: 5165

ABSTRACT

Objectives: The main aim of the study was to explore the potential sources of variation and understand the meaning of safety climate for nursing practice in acute hospital settings in the UK.

Design: A sequential mixed methods design included a cross-sectional survey using the Safety Climate Questionnaire (SCQ) and thematic analysis of focus group discussions. Confirmatory Factor Analysis (CFA) was used to validate the factor structure of the SCQ. Factor scores were compared between nurses working in operating theatres, critical care and ward areas. Results from the survey and the thematic analysis were then compared and synthesised.

Setting: A London University.

Participants: 319 registered nurses working in acute hospital settings completed the SCQ and a further 23 nurses participated in focus groups.

Results: CFA indicated that there was a good model fit on some criteria (χ^2 = 1683.699, df 824, $p < 0.001$; χ^2/ df = 2.04; RMSEA = 0.058) but a less acceptable fit on Comparative Fit Index (CFI) = 0.804. There was a statistically significant difference between clinical specialisms in Management Commitment (F [4,266] = 4.66, p = 0.001). Nurses working in operating theatres had lower scores compared with ward areas and they also reported negative perceptions about management in their focus group. There was significant variation in scores for Communication across clinical specialism (F [4,266] = 2.62, p = 0.035) but none of the pair-wise comparisons achieved statistical significance. Thematic analysis identified themes of Human Factors, Clinical Management and Protecting Patients. The System and the Human Side of Caring was identified as a meta-theme.

Conclusions: The results suggest that the SCQ has some utility but requires further exploration. The findings indicate that safety in nursing practice is a complex interaction between safety systems and the social and interpersonal aspects of clinical practice.

ARTICLE SUMMARY

Strengths and limitations of this study:

- The results of the study indicate that there is an important and complex link between human factor approaches used in nursing practice and the interpersonal aspects of care.
- This work makes a unique contribution to understanding safety climate in nursing practice in the UK setting.
- The Confirmatory Factor Analysis of the Safety Climate Questionnaire indicated that the model fit could be improved but further psychometric exploratory analysis may be warranted.
- The results need to be considered in the light of a cross-sectional survey response rate of 57% and a low number of participants in some of the focus groups.

The research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

INTRODUCTION

There is a growing consensus in healthcare safety research that organisational culture is critical for patient safety [1] and that safety management should move away from depending on lagging indicators of safety issues, such as incident reports, and move towards leading indicators, such as, measures of safety climate.[2] Patient safety culture is defined as aspects of organisational culture that are 'the product of individual and group values, attitudes, perceptions, competencies and patterns of behaviour that determine the commitment to and the style and proficiency of an organisations' health and safety management'.[3] Safety climate is defined as a measurable feature of staff's attitudes and perceptions of an organisations underlying safety culture at any point in time.[4] There is evidence that safety climate is open to change and has an impact on individual safety behaviour and an important factor in improving patient safety.[5,6]

The Safety Climate Questionnaire (SCQ) developed in the UK [7] has been used extensively in the NHS by the Royal College of Nursing.[8] However the SCQ was originally developed for use in the UK petroleum industry as part of a tool kit to measure safety climate. The SCQ measures nine factors that contribute to safety climate, namely Management commitment, Communication, Priority of safety, Safety rules and Procedures, Supportive environment, Involvement, Personal priorities and need for safety, Personal appreciation of risk and Work environment.[9] It is noted that the petroleum industry exhibits aspects of a High Reliability Organisation,[10] defined as 'organisations that are able to manage and sustain almost error-free performance despite operating in hazardous conditions where the consequences of errors could be catastrophic'[11] and as such lessons learnt from High Reliability Organisations have underpinned developments in safety and risk management in the NHS.[12] The petroleum industry is a very different setting from healthcare organisations but it is possible that their safety management systems could provide beneficial outcomes in safety and risk management in the healthcare setting.[13] Pilot testing of the SCQ undertaken within the NHS tested its usability and found that the tool was useable in this context.[14] However, neither an exploratory or confirmatory factor analysis of the tool was undertaken to validate its psychometric properties with a healthcare population.

Research evidence suggests that measures of safety climate vary between and within healthcare organisations and that there is limited understanding of the factors that may

influence and explain the sources of these variations.[15] Several research studies have reported safety climate scores varying across different clinical specialities with some reporting less safe climates in operating theatres, critical care and emergency departments compared to surgical and medical inpatient areas [16-18] and others reporting a safer climate in critical care [19-20]. However, none of this research has been undertaken in the UK. The underlying reasons for these variations in safety climate are unclear at the present time. Understanding the underlying factors that influence healthcare practitioner's perceptions of safety climate is important for the development of strategies to improve patient safety.

As a subset of healthcare practitioners, nurses make an important contribution to patient care and evidence indicates that nurse-staffing levels have a direct impact on patient mortality, [21-22] and nurse's perceptions of safety climate impacts upon safety behaviours and outcomes. [23] Therefore it is important to understand how nurses perceive safety climate as this may have a direct impact on patient safety. This mixed methods study set out to explore the underlying factors that contribute to safety climate in nursing practice. The main aim of the study was to explore the potential sources of variation in safety climate between different clinical specialities. The study set out to determine whether there are differences in the perception of safety climate between nurses working in critical care, operating theatres, surgical and medical wards in acute hospital settings in the UK and understand the meaning that nurses working in these different clinical settings attribute to their understanding of patient safety. The factor structure of the SCQ was also explored.

METHOD

The study design was a fully mixed, sequential, equal status, mixed methods design and was conducted in two phases.[24] The first phase of the study measured and then compared safety climate scores between groups of nurses working in operating theatres, critical care, surgical and medical ward areas. As the factor structure of the SCQ had not been evaluated in a nursing sample a Confirmatory Factor Analysis (CFA) was also undertaken. The results from the cross-sectional survey were used to structure the focus group discussions held with groups of nurses from operating theatres, critical care and ward areas. The results of both phases of the study were then jointly summarised in a statistics-by-theme format to facilitate

more in-depth inferences in order to consider potential mechanisms underlying safety climate.[25-26]

Following local ethical approval participants were recruited from a qualified nursing population who attended a university that recruited from a wide range of NHS Trusts and private hospitals in the region. In the UK Band 5 and 6 nurses are qualified nurses who deliver bedside care. They were specifically chosen, because they have a direct impact upon patient care and safety in their everyday practice. A convenience sampling method was used and participants were approached by the researcher at the beginning of a teaching session and the purpose of the survey was explained. Information sheets were included with the questionnaire and completion of the questionnaire implied consent. All questionnaires distributed were collected at the end of the afternoon teaching session. The aim was to collect at least 300 questionnaires as this is considered by some to be the minimum number required for robust factor analysis.[27]

A paper version of the SCQ was distributed to participants. Additional questions were added to the questionnaire in order to facilitate a stratified analysis to compare scores between nurses working in different clinical settings and measure potential factors that may influence perceptions of safety. These additional questions collected data on the clinical area the participant worked in, including whether they worked in a surgical ward, medical ward, critical care unit, operating theatre or other acute hospital unit. Further information included how long they had worked in their present position, how long they had worked in the speciality, how long they had been qualified and whether they had safety training and further training in their speciality. Participants were also asked to describe the type of training they had undertaken.

The SCQ has 43 questions with a 5-point Likert scale response and is scored by allocating a value of 5 to the ‘strongly agree’ response, 4 to ‘agree’ response, 3 to the ‘neither agree nor disagree’ response, 2 to the ‘disagree’ response and 1 to the ‘strongly disagree’ response. The negative worded questions were allocated a reverse score by subtracting the initial score from 6. The initial scores from the questionnaires provided raw scores and these were transferred

into an Excel® 2013 spreadsheet. In order to ensure that the data entry was as accurate as possible a double data entry procedure was followed as recommended by Elliot et al.[28]. The Excel spreadsheet was then transferred into SPSS® V21 and a Little's 'missing completely at random' (MCAR) test was undertaken to ensure that any missing data was not introducing bias into the analysis.[29]

A confirmatory factor analysis of the SCQ scores was undertaken using SPSS® Amos V21. The original nine factor structure as identified by Cox and Cheyne was used as the *a priori* model to be confirmed by the factor analysis.[7] The following goodness of fit indices were used to test the model. Chi Square (χ^2) and the χ^2/df ratio, the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The χ^2/df ratio overcomes the problem of a statistically significant χ^2 result associated with a larger sample sizes. A value of between 2 -3 is deemed as being acceptable the smaller value the better the fit.[30] The Comparative Fit Index (CFI) measures the difference in the non-centrality estimates of the baseline and proposed model with values ranging from 0 to 1. A cutoff value above 0.9 is considered to be an indication of a good model fit. The Root Mean Square Error of Approximation (RMSEA) measures the discrepancy between the hypothesised model and the population covariance matrices, and values range from 0 – 1. A RMSEA of less than 0.06 is indicative of an acceptable model fit with a recommended upper limit of 0.07.[31-32]

Once the CFA had been undertaken comparisons of safety climate dimensions (factors) were made between different clinical settings. Higher mean scores indicate a good safety climate. Dimension scores were compared between clinical specialisms using a general linear model (GLM) that adjusted for the following characteristics: years in current position, years qualified, years in specialism, specialist qualification and safety training. Adjusted means with 95%confidence intervals were calculated. Where there were differences between clinical specialism, based on the GLM F statistic, Bonferroni post-hoc pair wise comparisons were performed.

A Levene test of homogeneity of variance was conducted and residual plots produced, to ascertain whether the assumptions underpinning GLM had been met. A wild Bootstrap analysis was undertaken on the 'Personal priorities and need for safety' dimension to assess

whether non-equality of variance had biased the results. [33] The results remained very similar and only those from the GLM have been reported.

Following the survey a total of 23 nurses were recruited and participated in four focus groups (Operating theatre group = 8, Critical Care group = 9, ward A group = 3 and Ward B group = 3). A convenience sample method was used and participants were approached during a teaching session where information was provided and the purpose of the focus group was explained. The focus group discussions were arranged during lunch time. All participants consented to participate in the focus groups. These participants had not participated in the survey and therefore had not completed the SCQ. A priori open questions were used and participants were asked what their overall understanding of safety climate or culture was and what their views on communication and manager commitment to safety were, as these dimensions of safety climate had been found to be different between groups in the first phase of the study. Participants were not told the details of the differences found between different clinical settings in the survey. Each focus group was facilitated by one researcher who acted as facilitator, and an observer who noted group dynamics and timed the session. The groups lasted between 40 to 50 minutes and were recorded and later transcribed. A six phase approach to a thematic analysis was undertaken. [34] The transcribed discussions were imported into NVIVO 10 for windows to facilitate the development of codes. In-vivo coding was utilised for first order coding, as using the participants own words provided a much closer interpretation of their voice in the coding process. [35] The initial codes were refined throughout the process of analysis and codes were checked back to the transcripts to ensure that the meaning of the code was valid in the context of the content of the transcript. During second order coding, two researchers coded and the initial codes were reviewed and grouped into categories and eventually into sub-themes and themes. A process of checking coding between the researchers through discussion and agreement was undertaken to ensure reliability and validity of the coding process.

RESULTS

Survey results

A total of 563 questionnaires were distributed and 319 questionnaires were completed and returned (response rate = 57%). Four questionnaires were excluded from the final analysis because they were completed by nurses who did not fulfil the selection criteria, i.e., not a band 5 or 6 adult nurse working in an acute hospital setting. Little's MCAR indicated that the missing data were missing completely at random and were unlikely therefore to unduly affect the results (Little's MCAR test: $\chi^2 = 2368.11$, df = 2292, p = 0.131)

Table one illustrates the demographic data of the participants according to the specialist areas they worked in. There were more participants from critical care units than from other groups. The group identified as other included participants who stated that they worked in acute hospital setting areas such as, out patients, care of the elderly, oncology and haematology. The numbers of participants in these areas was low so these were grouped together.

Table One: Mean and standard deviation (SD) χ^2 and for demographic data for critical care, operating theatres, medicine, surgery and other clinical areas.

	Critical Care (n = 107)	Operating theatres (n = 49)	Medicine (n = 70)	Surgery (n= 54)	Other (n=24)	χ^2
Present position Mean (SD) years	3.12 (2.60)	4.26 (3.58)	3.69 (3.35)	3.12 (2.22)	3.21 (2.48)	p = 0.442
Years qualified Mean (SD) years	7.63 (5.47)	8.90 (6.85)	8.14 (6.46)	6.93 (6.00)	8.60 (6.04)	p = 0.317
Years Specialism Mean (SD) years	4.30 (3.77)	6.34 (5.25)	5.02 (3.84)	4.29 (3.74)	4.13 (2.70)	p = 0.195
Specialist qualification Percentage	50% (54/107)	43% (20/49)	37% (26/70)	33% (18/54)	58% (14/24)	P = 0.029*
Safety training Percentage	71% (76/107)	55% (27/49)	69% (48/70)	59% (32/54)	67% (16/24)	P = 0.032*

* statistically significant difference

Across the groups the participants had been working in their present position between 3 to 4 years. There was more variability across the groups in terms of how long the participants had been qualified with the critical care and surgery ward nurses being qualified as a registered nurse for less time. There was some variation in the amount of time the participants had been working within the specialism and the results indicate that the participants had been working in other areas before finally working within their specialist areas. The percentage of those reporting having undergone safety training ($\chi^2 = 6.12$, df = 4, p = 0.032) and those participants reporting having a specialist qualification ($\chi^2 = 9.83$, df = 4, p = 0.029) varied significantly across clinical specialism. All other variables did not vary significantly across clinical specialism. All participants who had reported undergoing safety training undertaken in UK hospitals on an annual basis described this as mandatory training. Typically this includes training in manual handling, resuscitation and infection control.

Confirmatory Factor Analysis

The CFA goodness of fit measures indicated that there was a good model fit on some criteria with a significant Chi Square test ($\chi^2 = 1687.560$, df = 824, p = < 0.001). Both the χ^2/df ratio of 2.05 and RMSEA value of 0.058 (90% CI interval 0.054 to 0.062) indicated a good model fit. However, the CFI was 0.805, although this was towards the higher end of the CFI range (0 to 1) it was below the acceptable threshold level (CFI >0.9) and suggests that the model could be improved.

The CFA regression weights (factor loadings) were similar to those from the original petroleum industry study (see supplementary table). However, there were four items that were particularly low and related to the dimensions of supportive environment, personal appreciation of risk and work environment. In relation to a supportive environment the item relating to, 'A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate', had a regression weight of 0.150 and the item relating to, 'When people ignore safety procedures here I feel it is none of my business', had a regression weight of 0.291. In the dimension of personal appreciation of risk, the item, 'I am rarely worried about being injured in the job', had a regression weight of 0.110 and in the dimension of Work environment the item, 'This is a safer place to work than other Trusts I have worked for', had a regression weight of 0.270. These items may not make a significant contribution to

the perception of safety climate in a nursing population. Cox and Cheyne [7] kept lower regression weighted items in their original questionnaire and suggested that these items should be used with caution.

Cronbach's alpha was greater than 0.70 for five of the nine dimensions (Management commitment 0.84, Priority of safety 0.76, Communication 0.70, Personal priorities and need for safety 0.72, Work environment 0.72). There were four dimensions with a Cronbach's alpha of less than 0.70 (Safety rules 0.67, Supportive environment 0.55, Involvement 0.58, Personal appreciation of risk 0.48). There was some marginal improvement in Cronbach's alpha when items with standardized regression weights of less than 0.3 were excluded (Supportive environment 0.55 to 0.57, Personal appreciation of risk 0.48 to 0.50, Work environment 0.72 to 0.74).

Comparison of safety climate scores

Following the CFA the factor scores derived from the survey were used to go onto explore differences in safety climate scores between nurses working in different clinical specialisms. Comparisons were made between nurses working in critical care areas, operating theatres, medical wards, surgical wards and other acute hospital settings as described above. Table two shows the adjusted GLM mean, 95% confidence interval by clinical specialism and F statistic, for all of the safety climate dimensions. Overall the scores were towards the higher range on the safety climate scale and suggested that participants reported a fairly positive safety climate for most of the dimensions. However, the work environment factor had lower scores across all the groups whilst personal priority of safety scored highly across all groups. There was a statistically significant difference between groups for Management Commitment ($F [4,266] = 4.66$, $p = 0.001$) and for Communication ($F [4,266] = 2.62$, $p = 0.035$).

Table Two: Comparison of the nine safety climate dimensions across clinical specialism adjusting for profile variable.

m = significantly different from Medicine; s= significantly different from surgery; o= significantly different from operating theatres

		Critical Care	Operating theatres	Medical wards	Surgical wards	Other	F test, p
Management Commitment	Mean	3.48	3.27[m,s]	3.75[o]	3.66[o]	3.31	F (4,266) = 4.66, p= 0.001
	(95% CI)	(3.34, 3.62)	(3.07, 3.67)	(3.59, 3.91)	(3.47, 3.85)	(2.99, 3.63)	
Priority of safety	Mean	3.54	3.44	3.73	3.50	3.61	F (4,266) = 1.29, p = 0.27
	(95% CI)	(3.39, 3.69)	(3.22, 3.66)	(3.55, 3.91)	(3.30, 3.71)	(3.26, 3.96)	
Communication	Mean	3.19	3.17	3.50	3.35	3.13	F (4,266) = 2.62, p= 0.035
	(95% CI)	(3.04, 3.33)	(2.96, 3.38)	(3.33, 3.67)	(3.15, 3.54)	(2.79, 3.47)	
Safety rules	Mean	3.18	3.23	3.43	3.40	2.90	F (4.266) = 1.96, p = 0.10
	(95% CI)	(3.01, 3.36)	(2.98, 3.48)	(3.22, 3.64)	(3.17, 3.64)	(2.49, 3.31)	
Supportive environment	Mean	3.66	3.67	3.86	3.75	3.63	F (4,266) = 1.85, p = 0.12
	(95% CI)	(3.55, 3.76)	(3.51, 3.82)	(3.73, 3.98)	(3.60, 3.89)	(3.38, 3.88)	
Involvement in safety	Mean	3.31	3.45	3.50	3.63	3.37	F (4,266) = 1.87, p = 0.12
	(95% CI)	(3.16, 3.46)	(3.24, 3.66)	(3.33, 3.68)	(3.43, 3.82)	(3.03, 3.71)	
Personal priorities and need for safety	Mean	4.20	4.31	4.37	4.33	4.11	F(4,266) = 1.89, p = 0.11
	(95% CI)	(4.10, 4.30)	(4.16, 4.45)	(4.25, 4.48)	(4.20, 447)	(3.88, 4.34)	
Personal appreciation of risk	Mean	3.19	3.15	3.36	3.44	3.35	F (4,226) = 0.92, p = .080
	(95% CI)	(3.05, 3.32)	(2.96, 3.34)	(3.20, 3.52)	(3.26, 3.61)	(3.04, 3.65)	
Work environment	Mean	2.62	2.65	2.68	2.82	2.85	F (4,266) = .092, p = 0.45
	(95% CI)	(2.47, 2.77)	(2.44, 2.86)	(2.50, 2.85)	(2.62, 3.02)	(2.51, 3.20)	

A Bonferroni post-hoc test revealed that there was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses (Mean = 3.27, 95% CI 3.07 – 3.47), compared with nurses working in medical wards (Mean = 3.75, 95% CI 3.59 – 3.91) and surgical ward settings (Mean = 3.66, 95% CI 3.47 – 3.85). Although there was significant variation in safety climate scores for communication across clinical specialism, none of the pair-wise comparisons achieved statistical significance at the 5% level, although the difference between critical care (Mean = 3.19, 95% CI 3.04 – 3.33) and the medical wards (Mean = 3.50 95% CI 3.33 – 3.67) came close (p = 0.056).

Thematic analysis

The results of the cross-sectional survey indicated a difference between nurses on the dimensions of Management commitment and though not statistically significant, Communication. During the focus groups participants were invited to discuss their understanding of safety culture and for their views of management and communication related to safety. Specific details of the differences found in the survey were not disclosed to the participants in order not to lead the discussion. Though these two aspects were discussed several other issues were also raised by participants. Three main themes emerged from the thematic analysis of the focus group data. These were Human Factors, Clinical Management and Protecting Patients. A further meta-theme was also identified as The System and Human Side of Caring.

Human Factors

The theme of Human Factors related to aspects of the environment such as design and staffing, the use of checklists and incident reporting. Aspects of physical environment were viewed as carrying potential risks and hazards to patients and the nurse is important in constantly checking equipment to ensure safety. For example, this participant stated that, *'I have to go round everywhere, checking the emergency crash call, check the monitors. The date they were serviced.'* (Critical Care group). Other participants recognised environmental design that has improved patient safety, such as, laminated flooring, '*We have a laminated grip flooring. They can still have a fall but it is much better for them.'* (Medical ward group).

The ratio of the numbers of patients to nurses was a concern, for example, *'Even in the current era, the ratio of nurses to patients is still a bit high. In terms of care, sometimes we are under so much pressure.'* (Medical Surgical group). All the groups mentioned the use of checklists. The operating theatre group mentioned the use of the World Health Organisation (WHO) checklist and the ward groups mentioned the use of intentional rounding. Though the content of these checklists are different they were seen as having advantages for patient safety and have been embedded in nursing practice. For example, '*We're very serious about protocols and policies as well and….we live by the checklist now*.' (Operating theatre group) and '*We have a checklist now and we check every single patient on the ward is safe.'* (Medical Surgical ward group).

There were ambivalent feelings regarding the use of incident reports where some participants viewed them as positive opportunities to learn from error, for example, '*You can learn from error, you can see it's not about blame culture.*' (Medical ward group), or were seen negatively, as this participant articulated, *'Yeah, a weapon, not something to help you. We're going to tell on you.*'(Critical care group). All of these approaches are systematic ways of managing error that are evident in nursing practice and the participants recognised the importance of these approaches to patient safety.

Clinical Management

The theme of Clinical Management related to communication processes and management behaviours that were relevant to the day-to-day management of patient care. Structured approaches to communication, such as handover, team briefing, and ward rounds were viewed as important for patient safety. Generally communication between nursing teams was seen as positive but communication between professions was identified as problematic, *'I think communication between nurses is good and between doctors and doctors is very good, but I think that there is a massive communication breakdown in people from different professions….I think information is lost all the time*.'(Critical care group) The role of the medical notes was viewed as being very important in communicating medical decisions to nursing staff but this was problematic for many participants. For example, *'Sometimes you are on night shift and you handover to the nurse who is taking over in the morning and you handover things that have happened and there's nothing written in the notes, nothing written*

by the doctors.' (Critical care group). The nurses perceived medical staff as not understanding the significance of the medical record for safe nursing care.

Manager behaviour was also identified as very important for the participants feeling supported in patient safety. Managers who were seen as approachable and proactive in managing patient safety were generally viewed as providing support for example, '*My manager tends to pay a lot of attention to those small details where the chart is not updated, he will remind staff, so he is very picky on the small things, which is good because it reminds everybody about what you are doing.*' (Medical Surgical group). Those managers who were seen as unsupportive tended to be reactive and not supportive of staff, for example, '*Just telling me what to do. It's just like another surgeon telling me what to do.'* (Operating theatre group).

Protecting patients

Protecting patients was a key theme that emerged as being important aspect of nursing practice relating to patient safety. This focused upon how nursing skill is applied to patient care and acting as a gatekeeper and advocate for patients. There was an overall sense that patients are vulnerable, for example, '*The nature of our patients we're receiving acutely unwell patients who are suffering from delirium and are vulnerable.*' (Medical ward group). There was a sense that nurses protect patients by ensuring safety whilst undertaking nursing tasks, for example, *'Administrating medication is a major thing and I think safety should be ensured all the time and I see we always check, because you've got a critically ill patient and the last thing you want is a drug error.'* (Critical care group). There was also a sense that nurses need to challenge others. For example, *'I think when it comes to patient safety everyone has to take responsibility for safety, the doctors just don't do it. We encourage, we try to make everyone to be attentive but you have to challenge them.'* (Critical care group). There was a clear sense that the participants felt that they had a role in protecting patients from harm.

Joint synthesis of survey and focus group findings

The results of the cross-sectional survey found a variation in the dimension of Communication between nurses working in critical care and medical wards, though pairwise

comparisons were not statistically significant. Table three shows the mean and 95% confidence intervals for the dimension of Communication and a summary of the themes identified in the thematic analysis of the focus group discussions. The ward focus groups identified nurse-to-nurse communication as important for patient safety and these groups had slightly higher safety climate scores in this area. The Critical Care and Operating theatre focus group highlighted challenges associated with nurse to doctor communication.

Table Three: Differences in the dimension of communication between critical care, operating theatres, medical and surgical wards for the SCQ and theme

SCQ Communication score	Summary of thematic analysis
Critical care Mean = 3.19 (3.04, 3,33)	The main mechanism for communication was the ward round. Problems were identified where communication was poor following a ward round or where medical staff do not record in the medical record.
Operating theatres Mean = 3.17 (2.96, 3.38)	The main mechanism for communication was the WHO checklist and team briefing. There were challenges associated with compliance with these approaches from surgeons.
Medical wards Mean = 3.50 (3.33, 3.67) Surgical wards Mean = 3.35 (3.15, 3.54)	The main focus of communication was related to handovers between nursing teams and ward rounds. These seem to work well.

There was a statistically significant difference in mean safety climate scores for management commitment between operating theatre nurses, compared with nurses working in medical and surgical ward settings, with operating theatres having a lower score for Management Commitment. Table four shows the mean and confidence intervals for the dimension of Management Commitment and the themes that were identified in the focus groups. The operating theatre group reported more reactive and unsupportive manager behaviours in the focus group discussion. Whereas, the other areas generally reported proactive and supportive

manager behaviours in the focus groups, the operating theatre focus group reported reactive style of management.

Table Four: Differences in the dimension of management commitment to patient safety, between critical care, operating theatres, medical and surgical wards for the SCQ and themes.

SCQ Management commitment score	Summary of thematic analysis
Critical care Mean = 3.48 (3.34, 3.62)	Being approachable and accessible to support staff. Having experience and clinical credibility.
Operating theatres Mean = 3.27 (3.07, 3.67)	The perception that manager take sides with medical staff, not providing help and advice to nurses when they approach managers for assistance, and having an agenda related to targets, managers side with the surgeons and do not support the nursing staff, that the rules do not apply to surgeons.
Medical wards Mean = 3.75 (3.39, 3.91) Surgical wards Mean = 3.66 (3.47, 3.85)	Being proactive in supporting patient safety and reminding staff about compliance to safety procedures. Working clinically in the area and having clinical credibility with the nursing staff was highly valued and being approachable and accessible to nursing staff when they feel that they need support with problems related to patient safety.

The System and Human Side of Caring

A meta-theme, or overarching theme was identified from the three main themes and was labelled, the system and the human side of caring. This holistic view of the data captures two aspects of patient safety that seemed to be apparent within the data. That is, the system in which caring takes place, and this includes the physical environment, the design of that environment, and the system processes that have been put in place to assist patient safety with

the use of checklists and incident reporting. These systematic organisational structures and processes provide the backdrop and the context in which caring takes place. The human side of caring includes the personal and the interpersonal aspects of care, the need to communicate within nursing teams and to handover care to each other. The relationship with clinical managers was important to provide support for safe clinical care. The importance of interaction with other disciplines and the problems associated with that was a key component. Finally, the acknowledgement of patient's vulnerability within the system, and that nurses feel it is an important aspect of their role to act as an advocate and to protect patients through acting as a gatekeeper. Safety lies within an interaction between these two aspects of the clinical environment.

DISCUSSION

The application of High Reliability Organisation theory has underpinned the approach to patient safety in the past decade in the UK.[12] and the introduction of Human Factor approaches to patient safety is high on the agenda in the UK at the present time. The results of this study indicates that though human factor approaches are an important aspect of safe nursing practice, these approaches need to be supported with communication and management behaviours that rely upon good interpersonal skills. The emergent meta-theme of the system and the human side of caring indicates that attitudes and organisational culture are shaped and developed within the context of the transpersonal and the results indicate that support and communication empower nurses to advocate and protect their patients. The advent and development of checklists, the implementation of human factor and high reliability approaches are important and these have had a significant impact on patient safety but this study highlights other aspects of social behaviour and communication that can have an impact on patient safety. Indeed, too much focus upon targets and processes can be counterproductive. [36]

The SCQ has been used in the NHS extensively, however, the factor structure had not validated within a healthcare population before its use. The results of the confirmatory factor analysis undertaken here with a nursing sample, indicated that the SCQ did have an acceptable level of model fit for some but not all criteria. The main focus of this study was to

explore and understand variation in safety climate between specialisms and the SCQ provided some measurement that enabled further exploration of this variation. However, further work needs to be undertaken to fully validate this tool in the healthcare context. This tool was used extensively in the NHS without confirmation of its factor structure and these results illustrate that it is important to ensure that tools developed in one context are evaluated for fit into another context.

The findings indicated there was a lower safety climate in operating theatres compared to ward areas for management commitment. Both critical care and operating theatre groups also scored lower for communication than medical ward areas, though this was close to, but not statistically different. This may seem surprising, given that in recent years there has been widespread introduction of High Reliability Organisation approaches into critical care units and operating theatres, such as the WHO checklist into operating theatres and the introduction of reliability and standardisation measures in intensive care units. [37-38]. However, these results are consistent with results from other countries and may indicate that there is a fundamental difference in safety climate in different clinical settings and it has been suggested that these differences are associated with the severity or complexity of the patient condition, high patient turnover, or the technological complexity of the care delivered.[16-18] The results of this mixed methods study may point to other factors associated with management and communication differences in these areas rather than the highly technical aspects of patient care associated with critical areas. It is interesting to note that the SCQ does not stipulate whether management commitment indicates middle or senior management. It was clear in the focus group discussions that nurses see their ward or unit manager as their manager. How nurses interpret these issues has implications for how safety climate scores can be interpreted.

In a post Francis Inquiry [36] era, nursing care in particular has had increasing scrutiny of its practice, and these results indicate that there is a focus on safety in clinical practice and this is reflected in the perceptions and attitudes of the nurses who participated in this study. The factor scores of Personal priorities and need for safety, were consistently high across all groups, suggesting that for the participants, safety is an important priority in patient care for these nurses and this was reflected in the focus group discussions. The factor scores for Work environment were consistently low across all groups and the focus group discussions

highlighted the availability of equipment, staffing, the resources and time available to undertake the work are important aspects of safety in nursing practice.

It is acknowledged that the results need to be considered in the light of a cross-sectional survey response rate of 57% and the fact that the number of participants in some of the focus groups was low. However, the response rate is similar to other work undertaken in the field and although there were low numbers in some focus groups robust data was generated. However, the results of this study raise some important issues relating to the underlying drivers of safety climate in nursing practice and the importance of using a mixed methodology to provide a deeper insight into the mechanisms driving safety climate in nursing practice. Using a mixed methodology enabled a much deeper investigation of potential factors driving safety climate. The utilisation of mixed methodology and a further investigation of manager behaviours are potentially fruitful areas for further investigations in patient safety climate.

ACKNOWLEDGEMENTS

To the RCN for granting permission to use their version of the SCQ. To the participants who gave their time to complete the SCQ.

Original data is held by the corresponding author but arrangements are being made for data to be deposited onto Dryad.

Contributorship Statement:

M Tarling made a substantial contribution to the initial concept and design, data collection and analysis, drafting and revising the work and was main editor for this submission.

A Jones made a substantial contribution to the development of the design, qualitative analysis, drafting and revision of the work and final approval of the published version.

T Murrells made a substantial contribution to the development of the design, analysis of the quantitative data, drafting and revision of the work and final approval of the published version.

H McCutcheon made a substantial contribution to the development of the design, qualitative data analysis, drafting and revising the work and final approval of the published version.

REFERENCES

1. The Health Foundation. Measuring Safety Culture. London: The Health Foundation 2011.
2. The Health Foundation. The measurement and monitoring of safety. London: The Health Foundation. 2013.
3. Health and Safety Commission (HSC). Organising for Safety: Third Report. ACSNI Study Group on Human Factors. Sudbury: HSE Books. 1993 P15.
4. Health and Safety Executive. Reducing error and influencing behaviour. London: HSE. 1999.
5. Pronovost P and Sexton B. Assessing safety culture: guidelines and recommendations. *Qual Saf Health Care*. 2005;14:231–233.
6. Freeth D, Sandall J, Allan T, et al. A methodological study to compare survey-based and observation-based evaluations of organisational and safety cultures and then compare both approaches with makers of the quality of care. Health Technology Assessment. 2012;6:25.
7. Cox SJ and Cheyne AJT. Assessing safety culture in offshore environments. *Saf Sci.* 2000;34:111–129.
8. Currie L and Watterson L. Measuring the safety climate in NHS organisations. *Nurs Stand*. 2010;24(24):35–38.
9. University of Loughborough. Safety climate measurement user guide and toolkit. 2000. Available at: http://www.lboro.ac.uk/departments/sbe/downloads/pmdc/safety-climate-assessment-toolkit.pdf (assessed July 2013)
10. The Health Foundation. High Reliability Organisations. London: The Health Foundation. 2011.
11. Health and Safety Executive. High reliability organisations: A review of the literature. London: HSE. 2011.
12. Department of Health. An Organisation with a memory. London: The Stationary Office. 2001.
13. Olsen E and Aase KA. Comparative study of safety climate differences in healthcare and the petroleum industry. *Qual Saf Health Care.* 2010;19(Suppl 3):175–179.
14. Currie L and Cooper C. Measuring staffs' perceptions of safety climate: testing The Safety Climate Questionnaire (SCQ) for use in healthcare. London: RCN. 2008.
15. Zohar D, Livne Y, Tenne-Gazit O, et al. Healthcare climate: a framework for measuring and improving patient safety. *Critical Care Medicine.* 2007;35:1312–1317.

16. Hartmann CW, Rosen AK, Meterfo M, et al. An overview of patient safety climate in the VA. *Health Services Research.* 2008;43(4):1263–1284.
17. Singer SJ, Gaba DM, Fulwell A, et al. Patient safety climate in 92 US hospital: differences by work area and discipline. *Medical Care.* 2009;47:23–31.
18. Campbell EG, Singer S, Kitch BT, et al. Patient Safety Climate in Hospitals: Act Locally on Variation Across Units. *Joint Commission on Accreditation of Healthcare Organizations.* 2010;36(7):316–326.
19. Abbas HAE, Bassiuni NA, Baddar FM Perception of front-line healthcare providers toward patient safety: A preliminary study in a university hospital in Egypt. Topics in Advanced Practice Nursing e-journal. 2008:8(2): Medscape.
20. Abdou HA and Saber KM A baseline assessment of patient safety culture among nurses at student university hospital. World Journal of Medical Sciences. 2011:6(1): 17 – 26.
21. Needleman J, Buerhaus P, Pankratz VS, et al. Nurse staffing and inpatient hospital mortality. *N Engl J Med.* 2011;364:1037-1045.
22. Aiken LH, Sermeus W, Van den Heede K et al Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. British Medical Journal. 2012:344 e1717.
23. Agnew C, Flin RH, Mearns KJ. (2013). Patient safety climate and worker safety behaviours in acute hospitals in Scotland. Journal of Safety Research. 2013: 45: 95-101.
24. Leech N and Onwuegbuzie AJ (2009) A typology of mixed methods research designs. Quality and Quantity. 43: 265 – 275
25. Glynos J and Howarth D (2007) Logics of Critical Explanation in Social and Political Theory. London: Routledge.
26. Guetterman TC, Fetters MD, Creswell JW Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med. 2015 Nov; 13(6): 554–561.
27. Comrey, AL and Lee, HB*A. First Course in Factor Analysis.* Hillsdale, NJ: Erlbaum. 1992.
28. Elliot AC, Hynan LS, Reisch S, et al. Preparing data for analysis using Microsoft Excel. *J Investig Med.* 2006;54:6:334–341.
29. Little RJA and Rubin DB. Statistical analysis with missing data (2nd ed.). New York: Wiley. 2002.
30. Kline P. An easy guide to factor analysis. London: Routledge. 2008.

31. Hu L and Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*.1999;6(1):1-55.
32. Steiger JH. Understanding the limitations of global fit assessment in structural equation modelling. *Individual Differences.* 2007;42:893–898.
33. Lui RY. Bootstrap procedures under some Non-I.I.D. models. The Annals of Statistics. 1988;4:1696-1708.
34. Braun V and Clarke V Using thematic analysis in psychology. Qualitative Research in Psychology. 2006.3 (2): 77-101.
35. Saldaňa J The coding manual for qualitative researchers. (2nd Ed). 2013. London: Sage Publications.
36. Francis R Report of the Mid-Staffordshire NHS Foundation Trust public inquiry. 2013 London: The Stationary Office.
37. Haynes AB, Weiser TG, Berry WR, et al. (2009) A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;360:491-499.
38. Pronovost PJ, Berenholtz SM, Goeschel CA. Creating High Reliability in Health Care Organizations. *Health Services Research.* 2006;41(4) Part II:1599–1617.

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Management Commitment			
Managers and supervisors express concern if safety procedures are not followed	0.440	0.633	0.696
In my workplace managers/supervisors show interest in my safety	0.520	0.689	0.727
In my workplace management turn a blind eye to safety issues	0.737	0.618	0.712
In my workplace management acts quickly to correct safety problems	0.811	0.743	0.806
Corrective action is always taken when management is told about unsafe practice	0.690	0.527	0.571
Management acts only after incidents have occurred	0.500	0.666	0.710
Management acts decisively when a safety concern is raised	0.792	0.668	0.732
Priority of Safety			
Management considers safety to be equally important as getting the work done	0.534	0.717	0.757
Safety rules and procedures are carefully followed	0.585	0.592	0.670
Management clearly considers the safety of staff of great importance	0.665	0.659	0.688
I believe that safety issues are not assigned a high priority	0.585	0.675	0.742
Communication			
There is good communication here about safety issues which affect me	0.731	0.596	0.653
Safety information is always brought to my attention by my line manager/supervisor	0.633	0.596	0.638
I do not receive praise for working safely	0.481	0.511	0.569
My line manager/supervisor does not always inform me of current concern and issues	0.594	0.636	0.703
Management operates an open door policy on safety issues	0.541	0.511	0.574
Safety rules			
Some safety rules and procedures do not need to be followed to get the job done safely	0.724	0.600	0.580
Some health and safety rules and procedures are not really practical	0.685	0.622	0.744
Sometimes it is necessary to depart from safety requirements in order to get the work done	0.583	0.699	0.709
Supportive environment			
I can influence health and safety here	0.543	0.460	0.479
A no blame approach is used to persuade people acting unsafely that their behaviour is inappropriate	0.367	0.150*	0.153*
I am strongly encouraged to report unsafe conditions	0.639	0.692	0.756
When people ignore safety procedures here I feel it is none of my business	0.480	0.291*	0.387*
Co-workers often give tips to each other on how to work safely	0.323	0.452	0.498
Employees are not encouraged to raise safety concerns	0.421	0.508	0.624

Table Supplementary: Standardized regression weights from petroleum industry data and nursing data comparing AMOS and MPLUS.
*indicates items with low regression weights

	Industry sample	Nursing sample AMOS	Nursing Sample MPLUS
Personal priorities and need for safety			
A safe place to work has a lot of personal meaning to me	0.571	0.519	0.600
It is important to me that there is a continuing emphasis on safety	0.655	0.560	0.625
I understand the safety rules for my job	0.642	0.664	0.832
Personally I feel that safety issues are not the most important aspect of my job	0.500	0.571	0.694
Safety is the number one priority in my mind when completing a job	0.623	0.617	0.685
Personal appreciation of risk			
I am clear about what my responsibilities are for health and safety	0.273	0.561	0.744
I am sure it is only a matter of time before I am involved in an incident	0.782	0.382	0.383
In my workplace the chance of being involved in an incident are high	0.464	0.548	0.570
I am rarely worried about being injured on the job	0.286	0.110*	0.086*
Involvement			
I am involved with safety issues at work	0.687	0.657	0.602
I am never involved in the ongoing review of safety	0.524	0.402	0.541
I am involved in informing management of important safety issues	0.724	0.671	0.677
Work environment			
There are always enough people available to get the job done safely	0.596	0.527	0.593
Sometimes I am not given enough time to get the job done safely	0.668	0.610	0.543
Sometimes conditions here hinder my ability to work safely	0.666	0.724	0.727
Operational targets often conflict with safety measures	0.795	0.590	0.662
I cannot always get the equipment I need to do the job safely	0.448	0.587	0.645
This is a safer place to work than other Trusts I have worked for	0.256	0.270*	0.392*

McGill

Mixed Methods Appraisal Tool (MMAT) – Version 2011

For dissemination, application, and feedback: Please contact pierre.pluye@mcgill.ca, Department of Family Medicine, McGill University, Canada.

The MMAT is comprised of two parts (see below): criteria (Part I) and tutorial (Part II). While the content validity and the reliability of the pilot version of the MMAT have been examined, this critical appraisal tool is still in development. Thus, the MMAT must be used with caution, and users' feedback is appreciated. Cite the present version as follows.

> Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'Cathain, A., Griffiths, F., Boardman, F., Gagnon, M.P., & Rousseau, M.C. (2011). *Proposal: A mixed methods appraisal tool for systematic mixed studies reviews*. Retrieved on [date] from http://mixedmethodsappraisaltoolpublic.pbworks.com. Archived by WebCite® at http://www.webcitation.org/5tTRTc9yJ

Purpose: The MMAT has been designed for the appraisal stage of complex systematic literature reviews that include qualitative, quantitative and mixed methods studies (mixed studies reviews). The MMAT permits to concomitantly appraise and describe the methodological quality for three methodological domains: mixed, qualitative and quantitative (subdivided into three sub-domains: randomized controlled, non-randomized, and descriptive). Therefore, using the MMAT requires experience or training in these domains. E.g., MMAT users may be helped by a colleague with specific expertise when needed. The MMAT allows the appraisal of most common types of study methodology and design. For appraising a qualitative study, use section 1 of the MMAT. For a quantitative study, use section 2 or 3 or 4, for randomized controlled, non-randomized, and descriptive studies, respectively. For a mixed methods study, use section 1 for appraising the qualitative component, the appropriate section for the quantitative component (2 or 3 or 4), and section 5 for the mixed methods component. For each relevant study selected for a systematic mixed studies review, the methodological quality can then be described using the corresponding criteria. This may lead to exclude studies with lowest quality from the synthesis, or to consider the quality of studies for contrasting their results (e.g., low quality vs. high).

Scoring metrics: For each retained study, an overall quality score may be not informative (in comparison to a descriptive summary using MMAT criteria), but might be calculated using the MMAT. Since there are only a few criteria for each domain, the score can be presented using descriptors such as *, **, ***, and ****. For qualitative and quantitative studies, this score can be the number of criteria met divided by four (scores varying from 25% (*) -one criterion met- to 100% (****) -all criteria met-). For mixed methods research studies, the premise is that the overall quality of a combination cannot exceed the quality of its weakest component. Thus, the overall quality score is the lowest score of the study components. The score is 25% (*) when *QUAL=1 or QUAN=1 or MM=0*; it is 50% (**) when *QUAL=2 or QUAN=2 or MM=1*; it is 75% (***) when *QUAL=3 or QUAN=3 or MM=2*; and it is 100% (****) when *QUAL=4 and QUAN=4 and MM=3* (QUAL being the score of the qualitative component; QUAN the score of the quantitative component; and MM the score of the mixed methods component).

Rationale: There are general criteria for planning, designing and reporting mixed methods research (Creswell and Plano Clark, 2010), but there is no consensus on key specific criteria for appraising the methodological quality of mixed methods studies (O'Cathain, Murphy and Nicholl, 2008). Based on a critical examination of 17 health-related systematic mixed studies reviews, an initial 15-criteria version of MMAT was proposed (Pluye, Gagnon, Griffiths and Johnson-Lafleur, 2009). This was pilot tested in 2009. Two raters assessed 29 studies using the pilot MMAT criteria and tutorial (Pace, Pluye, Bartlett, Macaulay et al., 2010). Based on this pilot exercise, it is anticipated that applying MMAT may take on average 15 minutes per study (hence efficient), and that the Intra-Class Correlation might be around 0.8 (hence reliable). The present 2011 revision is based on feedback from four workshops, and a comprehensive framework for assessing the quality of mixed methods research (O'Cathain, 2010).

Conclusion: The MMAT has been designed to appraise the *methodological quality* of the studies retained for a systematic mixed studies review, not the quality of their *reporting* (writing). This distinction is important, as good research may not be 'well' reported. If reviewers want to genuinely assess the former, companion papers and research reports should be collected when some criteria are not met, and authors of the corresponding publications should be contacted for additional information. Collecting additional data is usually necessary to appraise *qualitative research and mixed methods studies*, as there are no uniform standards for reporting study characteristics in these domains (www.equator-network.org), in contrast, e.g., to the CONSORT statement for reporting randomized controlled trials (www.consort-statement.org).

Authors and contributors: Pierre Pluye[1], Marie-Pierre Gagnon[2], Frances Griffiths[3] and Janique Johnson-Lafleur[1] proposed an initial version of MMAT criteria (Pluye et al., 2009). Romina Pace[1] and Pierre Pluye[1] led the pilot test. Gillian Bartlett[1], Belinda Nicolau[4], Robbyn Seller[1], Justin Jagosh[1], Jon Salsberg[1] and Ann Macaulay[1] contributed to the pilot work (Pace et al., 2010). Pierre Pluye[1], Émilie Robert[5], Margaret Cargo[6], Alicia O'Cathain[7], Frances Griffiths[3], Felicity Boardman[3], Marie-Pierre Gagnon[2], Gillian Bartlett[1], and Marie-Claude Rousseau[8] contributed to the present 2011 version.

Affiliations: 1. Department of Family Medicine, McGill University, Canada; 2. Faculté des sciences infirmières, Université Laval, Canada; 3. Warwick Medical School, University of Warwick, UK; 4. Faculty of Dentistry, McGill University, Canada; 5. Centre de recherche du CHUM, Université de Montréal, Canada; 6. School of Health Sciences, University of South Australia, Australia; 7. Medical Care Research Unit, ScHARR, University of Sheffield, UK; 8. INRS-Institut Armand Frappier, Laval, Canada.

PART I. MMAT criteria & one-page template (to be included in appraisal forms)

Types of mixed methods study components or primary studies	Methodological quality criteria (see tutorial for definitions and examples)	Responses			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	• Are there clear qualitative and quantitative research questions (or objectives*), or a clear mixed methods question (or objective*)?				Page 5
	• Do the collected data allow address the research question (objective)? E.g., consider whether the follow-up period is long enough for the outcome to occur (for longitudinal studies or study components).				Page 5
	Further appraisal may be not feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.				
1. Qualitative	1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?				Page 8
	1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?				Page 8
	1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?				Page 8 & Page 18
	1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?				Page 8
2. Quantitative randomized controlled (trials)	2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?				N/A
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)?				N/A
	2.3. Are there complete outcome data (80% or above)?				N/A
	2.4. Is there low withdrawal/drop-out (below 20%)?				N/A
3. Quantitative non-randomized	3.1. Are participants (organizations) recruited in a way that minimizes selection bias?				N/A
	3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?				N/A
	3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?				N/A
	3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?				N/A
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?				Page 6
	4.2. Is the sample representative of the population understudy?				Page 6
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)?				Page 6
	4.4. Is there an acceptable response rate (60% or above)?				Page 8 & page 18
5. Mixed methods	5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?				Page 6
	5.2. Is the integration of qualitative and quantitative data (or results*) relevant to address the research question (objective)?				Page 14 - 16
	5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results*) in a triangulation design?				Page 18
	Criteria for the qualitative component (1.1 to 1.4), and appropriate criteria for the quantitative component (2.1 to 2.4, or 3.1 to 3.4, or 4.1 to 4.4), must be also applied.				

*These two items are not considered as double-barreled items since in mixed methods research, (1) there may be research questions (quantitative research) or research objectives (qualitative research), and (2) data may be integrated, and/or qualitative findings and quantitative results can be integrated.

PART II. MMAT tutorial

Types of mixed methods study components or primary studies	Methodological quality criteria
1. Qualitative Common types of qualitative research methodology include: A. Ethnography The aim of the study is to describe and interpret the shared cultural behaviour of a group of individuals. B. Phenomenology The study focuses on the subjective experiences and interpretations of a phenomenon encountered by individuals. C. Narrative The study analyzes life experiences of an individual or a group. D. Grounded theory Generation of theory from data in the process of conducting research (data collection occurs first). E. Case study In-depth exploration and/or explanation of issues intrinsic to a particular case. A case can be anything from a decision-making process, to a person, an organization, or a country. F. Qualitative description There is no specific methodology, but a qualitative data collection and analysis, e.g., in-depth interviews or focus groups, and hybrid thematic analysis (inductive and deductive). Key references: Creswell, 1998; Schwandt, 2001; Sandelowski, 2010.	**1.1. Are the sources of qualitative data (archives, documents, informants, observations) relevant to address the research question (objective)?** E.g., consider whether (a) the selection of the participants is clear, and appropriate to collect relevant and rich data; and (b) reasons why certain potential participants chose not to participate are explained. **1.2. Is the process for analyzing qualitative data relevant to address the research question (objective)?** E.g., consider whether (a) the method of data collection is clear (in depth interviews and/or group interviews, and/or observations and/or documentary sources); (b) the form of the data is clear (tape recording, video material, and/or field notes for instance); (c) changes are explained when methods are altered during the study; and (d) the qualitative data analysis addresses the question. **1.3. Is appropriate consideration given to how findings relate to the context, e.g., the setting, in which the data were collected?*** E.g., consider whether the study context and how findings relate to the context or characteristics of the context are explained (how findings are influenced by or influence the context). "For example, a researcher wishing to observe care in an acute hospital around the clock may not be able to study more than one hospital. (…) Here, it is essential to take care to describe the context and particulars of the case [the hospital] and to flag up for the reader the similarities and differences between the case and other settings of the same type" (Mays & Pope, 1995). The notion of context may be conceived in different ways depending on the approach (methodology) tradition. **1.4. Is appropriate consideration given to how findings relate to researchers' influence, e.g., through their interactions with participants?*** E.g., consider whether (a) researchers critically explain how findings relate to their perspective, role, and interactions with participants (how the research process is influenced by or influences the researcher); (b) researcher's role is influential at all stages (formulation of a research question, data collection, data analysis and interpretation of findings); and (c) researchers explain their reaction to critical events that occurred during the study. The notion of reflexivity may be conceived in different ways depending on the approach (methodology) tradition. E.g., "at a minimum, researchers employing a generic approach [qualitative description] must explicitly identify their disciplinary affiliation, what brought them to the question, and the assumptions they make about the topic of interest" (Caelli, Ray & Mill, 2003, p. 5).

*See suggestion on the MMAT wiki homepage (under '2011 version'): Independent reviewers can establish a common understanding of these two items prior to beginning the critical appraisal.

Types of mixed methods study components or primary studies	Methodological quality criteria
2. Quantitative randomized controlled (trials) Randomized controlled clinical trial: A clinical study in which individual participants are allocated to intervention or control groups by randomization (intervention assigned by researchers). Key references: Higgins & Green, 2008; Porta, 2008; Oxford Center for Evidence based medicine, 2009.	**2.1. Is there a clear description of the randomization (or an appropriate sequence generation)?** In a randomized controlled trial, the allocation of a participant (or a data collection unit, e.g., a school) into the intervention or control group is based solely on chance, and researchers describe how the randomization schedule is generated. "A simple statement such as 'we randomly allocated' or 'using a randomized design' is insufficient". *Simple randomization:* Allocation of participants to groups by chance by following a predetermined plan/sequence. "Usually it is achieved by referring to a published list of random numbers, or to a list of random assignments generated by a computer". *Sequence generation:* "The rule for allocating interventions to participants must be specified, based on some chance (random) process". Researchers provide sufficient detail to allow a readers' appraisal of whether it produces comparable groups. E.g., blocked randomization (to ensure particular allocation ratios to the intervention groups), or stratified randomization (randomization performed separately within strata), or minimization (to make small groups closely similar with respect to several characteristics).
	2.2. Is there a clear description of the allocation concealment (or blinding when applicable)? *The allocation concealment protects assignment sequence until allocation.* E.g., researchers and participants are unaware of the assignment sequence up to the point of allocation. E.g., group assignment is concealed in opaque envelops until allocation. *The blinding protects assignment sequence after allocation.* E.g., researchers and/or participants are unaware of the group a participant is allocated to during the course of the study.
	2.3. Are there complete outcome data (80% or above)? E.g., almost all the participants contributed to almost all measures.
	2.4. Is there low withdrawal/drop-out (below 20%)? E.g., almost all the participants completed the study.

Types of mixed methods study components or primary studies	Methodological quality criteria
3. Quantitative non-randomized Common types of design include (A) non-randomized controlled trials, and (B-C-D) observational analytic study or component where the intervention/exposure is defined/assessed, but not assigned by researchers. A. Non-randomized controlled trials The intervention is assigned by researchers, but there is no randomization, e.g., a pseudo-randomization. A non-random method of allocation is not reliable in producing alone similar groups. B. Cohort study Subsets of a defined population are assessed as exposed, not exposed, or exposed at different degrees to factors of interest. Participants are followed over time to determine if an outcome occurs (prospective longitudinal). C. Case-control study Cases, e.g., patients, associated with a certain outcome are selected, alongside a corresponding group of controls. Data is collected on whether cases and controls were exposed to the factor under study (retrospective). D. Cross-sectional analytic study At one particular time, the relationship between health-related characteristics (outcome) and other factors (intervention/exposure) is examined. E.g., the frequency of outcomes is compared in different population sub-groups according to the presence/absence (or level) of the intervention/exposure. Key references for observational analytic studies: Higgins & Green, 2008; Wells, Shea, O'Connell, Peterson, et al., 2009.	**3.1. Are participants (organizations) recruited in a way that minimizes selection bias?** At recruitment stage: For cohort studies, e.g., consider whether the exposed (or with intervention) and non-exposed (or without intervention) groups are recruited from the same population. For case-control studies, e.g., consider whether same inclusion and exclusion criteria were applied to cases and controls, and whether recruitment was done independently of the intervention or exposure status. For cross-sectional analytic studies, e.g., consider whether the sample is representative of the population. **3.2. Are measurements appropriate (clear origin, or validity known, or standard instrument; and absence of contamination between groups when appropriate) regarding the exposure/intervention and outcomes?** At data collection stage: E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) the measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure. For non-randomized controlled trials, the intervention is assigned by researchers, and so consider whether there was absence/presence of a contamination. E.g., the control group may be indirectly exposed to the intervention through family or community relationships. **3.3. In the groups being compared (exposed vs. non-exposed; with intervention vs. without; cases vs. controls), are the participants comparable, or do researchers take into account (control for) the difference between these groups?** At data analysis stage: For cohort, case-control and cross-sectional, e.g., consider whether (a) the most important factors are taken into account in the analysis; (b) a table lists key demographic information comparing both groups, and there are no obvious dissimilarities between groups that may account for any differences in outcomes, or dissimilarities are taken into account in the analysis. **3.4. Are there complete outcome data (80% or above), and, when applicable, an acceptable response rate (60% or above), or an acceptable follow-up rate for cohort studies (depending on the duration of follow-up)?**

Types of mixed methods study components or primary studies	Methodological quality criteria
4. Quantitative descriptive studies Common types of design include single-group studies: A. Incidence or prevalence study without comparison group In a defined population at one particular time, what is happening in a population, e.g., frequencies of factors (importance of problems), is described (portrayed). B. Case series A collection of individuals with similar characteristics are used to describe an outcome. C. Case report An individual or a group with a unique/unusual outcome is described in details. Key references: Critical Appraisal Skills Programme, 2009; Draugalis, Coons & Plaza, 2008.	**4.1. Is the sampling strategy relevant to address the quantitative research question (quantitative aspect of the mixed methods question)?** E.g., consider whether (a) the source of sample is relevant to the population under study; (b) when appropriate, there is a standard procedure for sampling, and the sample size is justified (using power calculation for instance).
	4.2. Is the sample representative of the population understudy? E.g., consider whether (a) inclusion and exclusion criteria are explained; and (b) reasons why certain eligible individuals chose not to participate are explained.
	4.3. Are measurements appropriate (clear origin, or validity known, or standard instrument)? E.g., consider whether (a) the variables are clearly defined and accurately measured; (b) measurements are justified and appropriate for answering the research question; and (c) the measurements reflect what they are supposed to measure.
	4.4. Is there an acceptable response rate (60% or above)? The response rate is not pertinent for case series and case report. E.g., there is no expectation that a case series would include all patients in a similar situation.

Types of mixed methods study components or primary studies	Methodological quality criteria
5. Mixed methods Common types of design include: A. Sequential explanatory design The quantitative component is followed by the qualitative. The purpose is to explain quantitative results using qualitative findings. E.g., the quantitative results guide the selection of qualitative data sources and data collection, and the qualitative findings contribute to the interpretation of quantitative results. B. Sequential exploratory design The qualitative component is followed by the quantitative. The purpose is to explore, develop and test an instrument (or taxonomy), or a conceptual framework (or theoretical model). E.g., the qualitative findings inform the quantitative data collection, and the quantitative results allow a generalization of the qualitative findings. C. Triangulation design The qualitative and quantitative components are concomitant. The purpose is to examine the same phenomenon by interpreting qualitative and quantitative results (bringing data analysis together at the interpretation stage), or by integrating qualitative and quantitative datasets (e.g., data on same cases), or by transforming data (e.g., quantization of qualitative data). D. Embedded design The qualitative and quantitative components are concomitant. The purpose is to support a qualitative study with a quantitative sub-study (measures), or to better understand a specific issue of a quantitative study using a qualitative sub-study, e.g., the efficacy or the implementation of an intervention based on the views of participants. Key references: Creswell & Plano Clark, 2007; O'Cathain, 2010.	**5.1. Is the mixed methods research design relevant to address the qualitative and quantitative research questions (or objectives), or the qualitative and quantitative aspects of the mixed methods question (or objective)?** E.g., the rationale for integrating qualitative and quantitative methods to answer the research question is explained. **5.2. Is the integration of qualitative and quantitative data (or results) relevant to address the research question (objective)?** E.g., there is evidence that data gathered by both research methods was brought together to form a complete picture, and answer the research question; authors explain when integration occurred (during the data collection-analysis or/and during the interpretation of qualitative and quantitative results); they explain how integration occurred and who participated in this integration. **5.3. Is appropriate consideration given to the limitations associated with this integration, e.g., the divergence of qualitative and quantitative data (or results)?**

References

- Caelli, K., Ray, L., & Mill, J. (2003). 'Clear as Mud': Toward greater clarity in generic qualitative research. *International Journal of Qualitative Methods, 2(*2), 1-23.
- Creswell, J., & Plano Clark, V. (2007). *Designing and conducting mixed methods research*. London: Sage.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks: Sage.
- Critical Appraisal Skills Programme (2009). CASP appraisal tools. Retrieved on August 26, 2009 from: *www.phru.nhs.uk/pages/PHD/resources.htm*
- Draugalis, J.R., Coons, S.J., & Plaza, C.M. (2008). Best practices for survey research reports: a synopsis for authors and reviewers. *American Journal of Pharmaceutical Education, 72(1),* e11.
- Higgins, J.P.T. & Green, S. (2008). *Cochrane Handbook for Systematic Reviews of Interventions* - Version 5.0.1 [updated September 2008]. The Cochrane Collaboration. Retrieved on August 26, 2009 from www.cochrane-handbook.org
- Mays, N., & Pope, C. (1995). Qualitative Research: Rigour and qualitative research. *British Medical Journal, 311(6997),* 109-112.
- O'Cathain, A., Murphy, E. & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *Journal of Health Services Research and Policy, 13(2),* 92-98.
- O'Cathain, A. (2010). Assessing the quality of mixed methods research: Towards a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research (2nd edition)* (pp. 531-555). Thousand Oaks: Sage.
- Pace, R., Pluye, P., Bartlett, G., Macaulay, A., Salsberg, J., Jagosh, J., & Seller, R. (2010). *Reliability of a tool for concomitantly appraising the methodological quality of qualitative, quantitative and mixed methods research: a pilot study*. 38th Annual Meeting of the North American Primary Care Research Group (NAPCRG), Seattle, USA.
- Pluye, P., Gagnon, M.P., Griffiths, F. & Johnson-Lafleur, J. (2009). A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies, 46(4),* 529-46.
- Oxford Center for Evidence Based Medicine (2009). Levels of evidence. Retrieved on July 7, 2009 from www.cebm.net/levels_of_evidence.asp
- Porta, M. (2008). *A Dictionary of Epidemiology*. New York: Oxford University Press.
- Sandelowski, M. (2010). What's in a name? Qualitative description revisited. *Research in Nursing and Health, 33*(1), 77-84.
- Schwandt, T. (2001). *Dictionary of qualitive inquiry*. Thousand Oaks: Sage.
- Wells, G.A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2009). *The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses*. The Cochrane Non-Randomized Studies Method Group. Retrieved on July 7, 2009 from www.ohri.ca/programs/clinical_epidemiology/oxford.htm