The Safety Organizing Scale

Development and Validation of a Behavioral Measure of Safety Culture in Hospital Nursing Units

Timothy J. Vogus, PhD, and Kathleen M. Sutcliffe, PhD†*

Background: Evidence that medical error is a systemic problem requiring systemic solutions continues to expand. Developing a "safety culture" is one potential strategy toward improving patient safety. A reliable and valid self-report measure of safety culture is needed that is both grounded in concrete behaviors and is positively related to patient safety.

Objective: We sought to develop and test a self-report measure of safety organizing that captures the behaviors theorized to underlie a safety culture and demonstrates use for potentially improving patient safety as evidenced by fewer reported medication errors and patient falls.

Subjects: A total of 1685 registered nurses from 125 nursing units in 13 hospitals in California, Indiana, Iowa, Maryland, Michigan, and Ohio completed questionnaires between December 2003 and June 2004.

Research Design: The authors conducted a cross-sectional assessment of factor structure, dimensionality, and construct validity.

Results: The Safety Organizing Scale (SOS), a 9-item unidimensional measure of self-reported behaviors enabling a safety culture, was found to have high internal reliability and reflect theoretically derived and empirically observed content domains. The measure was shown to discriminate between related concepts like organizational commitment and trust, vary significantly within hospitals, and was negatively associated with reported medication errors and patient falls in the subsequent 6-month period.

Conclusions: The SOS not only provides meaningful, behavioral insight into the enactment of a safety culture, but because of the association between SOS scores and reported medication errors and patient falls, it also provides information that may be useful to registered nurses, nurse managers, hospital administrators, and governmental agencies.

From the *Department of Management and Organization Studies, Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee; and the †Department of Management and Organizations, Ross School of Business, University of Michigan, Ann Arbor, Michigan.

Supported by the Blue Cross Blue Shield of Michigan Foundation and the Rackham Graduate School of the University of Michigan.

Reprints: Timothy J. Vogus, PhD, 401 21st Avenue South, Nashville, TN 37203. E-mail: timothy.vogus@owen.vanderbilt.edu.

Copyright © 2006 by Lippincott Williams & Wilkins

ISSN: 0025-7079/07/4501-0046

Key Words: high reliability, medical error, patient safety, safety culture, safety organizing

(*Med Care* 2007;45: 46–54)

On the heels of several troubling reports by the Institute of Medicine regarding the incidence and consequences of medical error[1,2]—the failure of a medical or nursing action to be completed as intended (eg, patients falling from their beds because nurses do not pull up the bedside railing) or the use of a wrong plan to achieve an aim (eg, administering a wrong drug or dosage to a patient) that may result in harm—patient safety has begun to receive renewed attention by researchers and practitioners alike. Although the actual number of medical errors[3,4] and their preventability[5] have been intensely debated, the fact remains that such errors are frequent[6,7] and costly both in human lives and expense to the healthcare system.[8] In the search for ways to reduce the incidence of such errors and improve safety, recent research has concluded that the most fruitful solutions focus on defects in the system of care delivery that give rise to errors.[1,9] One such systemic solution frequently recommended to healthcare organizations is developing a "culture of safety." A safety culture is the product of the shared values, attitudes, and patterns of behavior that determine the observable degree of effort with which all organizational members direct their attention and actions toward minimizing patient harm that may result from the process of care delivery.[10,11] An effective safety culture relies on ongoing collecting, analyzing, and disseminating of information from errors as well as proactive checks on the organization's vital signs.[10]

Recent studies provide evidence that safety cultures and patient safety are enabled by supportive leaders.[12,13] An emerging body of literature has also begun to empirically link safety culture to medical errors[14,15] and mortality rates.[16] Although these studies are instructive, they offer only indirect evidence regarding *how* a safety culture can reduce the incidence of medical errors. Without a way to directly assess the behaviors underlying a safety culture poses problems because it impedes ascertaining the exact mechanisms through which leader and other interventions reduce errors. As a result, interventions cannot be designed to ensure that they trigger behaviors that reduce the incidence of error.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 1. RN Characteristics, Unit Characteristics, Hospital Characteristics, and Outcomes by Site

	Hospital 1	Hospital 2	Hospital 3	Hospital 4	Hospital 5	Hospital 6	Hospital 7	Hospital 8	Hospital 9	Hospital 10	Hospital 11	Hospital 12	Hospital 13
RN age*	39.2	41.91	42.95	38.09	38.89	42.26	38.81	42.6	40.7	42.59	40.66	40.03	44.4
Percent RNs with BSN	50	18.68	25.58	40.63	47.4	21.74	48.28	41.47	38.92	48.71	30.52	55.75	32.15
RN tenure (unit)*	9.4	9.39	6.48	5.34	7.10	7.45	5.41	4.78	5.96	6.76	8.26	6.69	5.91
RN tenure (hospital)*	10	13.55	8.44	8.82	10.83	12.13	10.01	10.6	9.73	11.53	12.59	8.76	10.1
RN tenure (nursing)*	13.15	16.89	15.87	13.38	15.12	15.64	14.17	16.05	14.31	16.04	16.02	13.53	14.33
Average patients per RN	6.5	4.19	4.04	5	4.87	4.96	2.33	4.25	3.93	4.36	3.11	4.96	4.17
No. units†	1/0	9/9	5/5	2/2	25/22	12/8	3/0	8/0	9/9	15/12	18/13	12/8	6/6
Average unit size‡	35	17	15.4	27.5	22.19	25.7	12.67	24.5	19.75	23.07	18	24.42	16.4
Reported medication errors	N/A	70	62	42	208	137	N/A	N/A	149	173	173	37	17
Reported falls	N/A	50	18	23	151	132	N/A	N/A	95	90	63	103	69
Response rate	50%	64.35%	47.35%	56.79%	42.5%	64.09%	61.66%	56.97%	43.03%	52.97%	59.1%	42.69%	47.79%
Location	Suburban	Rural	Rural	Suburban	Urban	Suburban	Urban	Urban	Suburban	Suburban	Rural	Urban	Urban

*Values for age and the 3 types of tenure are in years.
†The number before the slash is the number of units with survey respondents; the number after the slash is the number of units also providing data on reported medication errors and patient falls.
‡Average unit size is given in number of average number of beds per unit.
RN indicates registered nurse; Percent RNs with BSN, percentage of RNs on within the given hospital with a Bachelor of Science degree in Nursing; N/A, not available.

To better understand the behaviors underlying an effective safety culture, we have developed the Safety Organizing Scale (SOS). Our primary objective was to develop a measure that is both substantively and methodologically consistent with the concept of safety culture. This means that it must be behavioral, shared at the unit level (indicative of collective and shared processes rather than individual behaviors), and associated with indicators of patient safety. Our measure needs to be shared and grounded in concrete behaviors because if a safety culture is present and consequential, it must be manifest in both shared attitudes and observable behaviors.[19,20] We focus on safety culture at the level of the caregiving unit because a safety culture is enacted collectively with one's most proximate colleagues and because variations in safety cultures, attitudes, and outcomes are evident at the unit level even within the same hospital.[14,19,20]

Theoretically, we ground our development of the SOS in the detailed case studies of "high-reliability organizations" (HROs). HROs such as aircraft carrier flight decks[21] and nuclear power plants[22] are organizations that operate hazardous technologies in a nearly error-free manner under trying conditions rife with complexity, interdependence, and time pressure. Case studies of HROs as well as healthcare organizations following the principles of HROs[23] suggest that safety cultures enable their reliable performance. Safety culture is seen as coming to life in HROs through behavioral processes of "collective mindfulness" enacted by front-line employees.[23–26] Collective mindfulness consists of 5 interrelated behavioral processes: preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise.[24–26]

We validate the SOS using a sample of registered nurses (RNs) on hospital nursing units. We use RNs because they are the surveillance system for early detection of emerging errors and complications in care and are well-positioned to take action to minimize negative outcomes for patients,[27] including intercepting medication errors.[9] We also focus on 2 indicators of patient safety that are primarily influenced by RNs: reported medication errors and patient falls.

METHODS

Setting

The units and RNs for this study were drawn from private, nonprofit Catholic hospitals that are members of a large Catholic Health System in the United States with member hospitals in California, Idaho, Indiana, Iowa, Maryland, Michigan, and Ohio. We conducted our multisite cross-sectional study between December 2003 and June 2004 using a convenience sample of 13 hospitals from this system. The participating hospitals included 5 urban hospitals (in the Columbus, Ohio, Detroit, Michigan, and Washington DC, Metropolitan Statistical Areas), 5 midsized metropolitan hospitals (in the Fresno, California, South Bend, Indiana, Battle Creek/Kalamazoo, Michigan, and Grand Rapids/Muskegon, Michigan, Metropolitan Statistical Areas), and 3 rural hospitals (in the rural Michigan and rural Iowa Metropolitan Statistical Areas). In addition to varying in geographic location, the participating hospitals also varied in size from 89 beds to 478 acute care beds. To increase generalizability of the SOS within hospitals, we sampled and received responses from a wide array of inpatient units (n = 125). Additional details regarding hospital, unit, and RN characteristics are presented in Table 1.

Response Rate

We surveyed all RNs in the participating hospitals and received responses to 1685 of 3298 questionnaires sent for a

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 2. Correspondence Between SOS Survey Items and Processes of Collective Mindfulness

Concept[24–26]	Definition[24–26]	SOS Survey Item(s)
Preoccupation with failure	Operating with a chronic wariness of the possibility of unexpected events that may jeopardize safety by engaging in proactive and preemptive analysis and discussion	When giving report to an oncoming nurse, we usually discuss what to look out for We spend time identifying activities we do not want to go wrong
Reluctance to simplify interpretations	Taking deliberate steps to question assumptions and received wisdom to create a more complete and nuanced picture of ongoing operations	We discuss alternatives as to how to go about our normal work activities
Sensitivity to operations	Ongoing interaction and information-sharing about the human and organizational factors that determine the safety of a system as a whole	We have a good "map" of each other's talents and skills We discuss our unique skills with each other so we know who on the unit has relevant specialized skills and knowledge
Commitment to resilience	Developing capabilities to detect, contain, and bounce back from errors that have already occurred, but before they worsen and cause more serious harm	We talk about mistakes and ways to learn from them When errors happen, we discuss how we could have prevented them
Deference to expertise	During high-tempo times (ie, when attempting to resolve a problem or crisis), decision-making authority migrates to the person or people with the most expertise with the problem at hand, regardless of their rank	When attempting to resolve a problem, we take advantage of the unique skills of our colleagues When a patient crisis occurs, we rapidly pool our collective expertise to attempt to resolve it

SOS indicates Safety Organizing Scale.

response rate of 51.1%. Although this rate is consistent with published studies in the medical literature using a mailed questionnaire,[28] such a response rate may be problematic if it biases or meaningfully affects the conclusions. Therefore, we investigated whether the response rate for a given unit was correlated with the SOS as well as other variables of interest, including reported medication errors, patient falls, RN staffing levels, or unit size. None of the correlations between response rate and these variables were statistically significant. To further assess whether our sample was in any way biased, we conducted a series of *t* tests comparing the unit-level means reported by RN respondents with unit-level means provided by nurse managers derived from the hospital's human resource information systems to determine if the RNs that responded to our survey were significantly different from those who did not respond in terms of age, 3 types of tenure (within the nursing profession, with the hospital, and on the unit), and level of education. None of these *t* tests revealed significant differences between respondents and nonrespondents, which suggests response bias is less likely to threaten the validity of our analyses.

Measure Validation Strategy

An effective measure should be theory-based, reliable, valid, relevant to the unit of analysis, and relatively easy to administer. Therefore, our first step was to determine the content validity of the SOS—the degree to which an instrument captures its intended construct in a manner consistent with relevant theory. To do so, we developed survey items that were closely linked to the theoretical literature on HROs that identifies safe performance as being a function of "collective mindfulness" as reflected by the 5 processes of preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise.[24–26] Each of these processes is precisely defined and linked to corresponding survey items in Table 2. Additional descriptive data on the SOS and its items are provided in the Appendix and mean

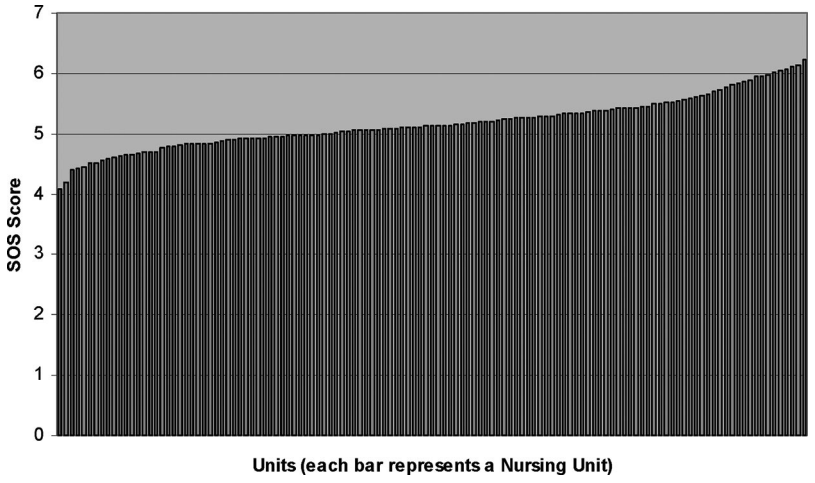

FIGURE 1. Mean levels of the Safety Organizing Scale by hospital nursing unit.

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

levels of the SOS by unit are displayed in Figure 1. We tailored our items to hospital nursing based on case studies of hospitals attempting to become HROs[23] and our own qualitative fieldwork in 2 hospital nursing units in a large Midwestern teaching hospital and in 2 geographically separate academic emergency departments. To further ensure that our survey items were consistent with existing theory and clear to RNs, we also had 7 experts in HROs and nursing assess the measure's content validity and pretested the SOS questionnaire with a sample of 45 RNs on a coronary care unit in a large Midwestern teaching hospital. Based on these pretests, we made one minor wording modification to a single item.

A measure is reliable when a set of items form an internally consistent scale. Reliability is a necessary but not sufficient condition for validity.[29] The construct validity of survey-based measures like the SOS is empirically determined and evidence of validity exists when a scale actually measures what it intends to measure (convergent validity), differentiates from other constructs (discriminant validity), and is related to its theoretical causes, correlates, and effects (criterion validity). It is important to note that we assess criterion validity using responses to the SOS aggregated to the unit level. We do so because the SOS is only theoretically meaningful at the unit level, and we want to assess the relationship between the SOS of a given nursing unit and its safety performance. We justify the use of the aggregate measure and conduct all other analyses related to convergent, discriminant, and criterion validity next.

RESULTS

Reliability and Convergent Validity

Convergent validity is the degree to which multiple items agree and is tested using confirmatory factor analysis (CFA). In this case, a CFA model estimated using AMOS 5.0[30] tests the hypothesis that the 9 items of the SOS are reflective of a single underlying factor. A CFA measurement model displays convergent validity if items load significantly onto the intended factor and model fit indices suggest adequate fit. The results of the measurement model are shown in Table 3, column 1.[31–33] Our model demonstrated excellent fit across all fit indices (CFI = 0.964, incremental fit index = 0.964, root mean square error of approximation = 0.055, standardized root mean square residual = 0.033) and all factor loadings were highly significant ($P < 0.001$). Taken in total, these results provide strong evidence for convergent validity. The Cronbach's alpha for the SOS was 0.88, strongly supporting scale reliability.

Discriminant Validity

We assessed discriminant validity by conducting "pairwise tests" of theoretically related constructs.[34] If items from other constructs are distinct from the SOS (ie, load on different factors), it is evidence of discriminant validity.[29] Two constructs theorized to be essential to the development of safety cultures—employee commitment and trust in manager[10,17]—served as comparison constructs. Commitment was measured using 3 items adapted from Meyer and Allen (eg, "I feel emotionally attached to my current work unit")[35] and 2 items for trust in manager ("My manager has a reputation for fairness in dealing with RNs" and "My manager demonstrates absolute integrity").

We evaluated discriminant validity by comparing the fit of 4 CFA models. First, we specified a 3-factor measurement model, including the SOS, commitment, and trust in manager. Second, we specified a 2-factor model in which one factor consisted of the items for the SOS and commitment and the second factor consisted of the items comprising trust in manager. Third, we specified a 2-factor model in which one factor consisted of the items for the SOS and trust in manager and the second factor consisted of the items comprising commitment. Lastly, we specified a one-factor model, including all the items for the SOS, commitment, and trust in manager scales. If the fit of the 3-factor model provides a significantly better fit than the alternative 2-factor (in which the SOS and commitment or the SOS and trust in manager load onto the same factor) or one-factor models using a χ^2 statistic, one can conclude that commitment and trust in manager are distinct from the SOS. Table 3 columns 3, 4, and 5 reports the results of these analyses and suggest that commitment and trust in manager are indeed distinct from the SOS (changes in χ^2 of 1193.17, 1670.99, and 2779.90, respectively, all differences significant at $P < 0.001$).

Statistical Justification for Aggregation

To ensure that the SOS was meaningful at the unit level, we constructed survey items to direct respondents' attention to common experiences of all RNs on the unit (eg, "we" or "RNs on this unit"). The survey instructions also encouraged respondents to assume the shared perspective of the group. Having items and directions that focus respondent attention on shared experience do, in fact, engender less within-group variability and more between-group variability than comparable survey items that reference individual experiences and perceptions.[36]

For aggregation to be statistically appropriate, it is necessary to demonstrate that 1) the members of each unit reported similar scores for the unit on a given measure and 2) the units have significant between unit variance for a given measure. Four complementary measures of within-group agreement were used to determine the degree of congruence between individual RNs' survey responses and the appropriateness of aggregating these measures to the unit level: the median $r_{wg(j)}$, the F-statistic from a one-way analysis of variance (ANOVA), intraclass correlation coefficient (ICC)(1) and ICC(2). $R_{wg(j)}$ measures the degree to which individual responses within a group are interchangeable; median $r_{wg(j)}$ values of 0.70 or greater provide evidence of acceptable agreement among individual responses on a scale.[37] Every unit in the sample had an $r_{wg(j)}$ value greater than 0.87 with a median value of 0.98. A significant F-statistic resulting from a one-way ANOVA with unit membership as the independent variable and the SOS as the dependent variable indicates that responses differ between RNs in different nursing units. The one-way ANOVA of

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 3. Confirmatory Factor Analyses Results for Convergent and Discriminant Validity of the SOS

Construct and Items	3-Factor Model (n = 1685)	2-Factor Model SOS/ Commitment and Trust in Manager (n = 1685)	2-Factor Model SOS/ Trust in Manager and Commitment (n = 1685)	1-Factor Model (n = 1685)
SOS				
Q1: Good "map" of skills.	0.606	0.616	0.599	0.607
Q2: Discuss mistakes to learn	0.723	0.705	0.711	0.694
Q3: Know others' skills	0.697	0.686	0.685	0.674
Q4: Discuss alternatives to normal work process	0.676	0.668	0.675	0.666
Q5: During report, discuss risks	0.559	0.540	0.550	0.532
Q6: Take advantage unique skills	0.670	0.663	0.665	0.657
Q7: Identify activities do not want to go wrong	0.683	0.666	0.680	0.662
Q8: Discuss error prevention	0.723	0.706	0.722	0.704
Q9: Pool expertise in crisis	0.613	0.607	0.613	0.604
Commitment				
Q1: Personal meaning in work unit	0.815	0.515	0.814	0.525
Q2: Emotionally attached to unit	0.802	0.505	0.805	0.513
Q3: Belonging in work unit	0.698	0.475	0.697	0.485
Trust in manager				
Q1: Reputation of fairness	0.889	0.890	0.464	0.481
Q2: Absolute integrity	0.931	0.930	0.480	0.497
Model fit				
χ^2 baseline model	10362.39 (df = 91)	10362.39 (df = 91)	10362.39 (df = 91)	10362.39 (df = 91)
χ^2 default model	442.24 (df = 72)	1635.41 (df = 74)	2113.23 (df = 74)	3222.14 (df = 75)
$\Delta\ \chi^2$ (from 3 factor)		1193.17 (df = 2)	1670.99 (df = 2)	2779.90 (df = 3)
CFI*	0.964	0.848	0.801	0.694
IFI†	0.964	0.848	0.802	0.694
RMSEA‡	0.055	0.112	0.128	0.158
SRMR§	0.033	0.075	0.073	0.095

Q1 through Q9 indicate the question number associated with the corresponding scale.

In all models, the error terms of Q2 and Q8 and Q6 and Q9 of the SOS are correlated.

*The comparative fit index[31] represents the fit of the model relative to the baseline model of independence among the observed variables. Values range from 0 to 1 with a good model fit being 0.95.[32]

†The incremental fit index[33] represents the fit of the model relative to the baseline model of independence among the observed variables. Values range from 0 to 1 with a good model fit being 0.95.[32]

‡The root mean square error of approximation is a population discrepancy function of the acceptability of the model that takes into account the complexity of the model. RMSEA assesses how well the model would fit the population covariance matrix were it known. A value less than 0.06 indicates good model fit.[32]

§The standardized root mean square residual is a summary measure of average standardized residuals when the model variance–covariance matrix is compared with the sample data variance–covariance matrix. Values range from 0 to 1. Values less than 0.08 represent good model fit.[32]

SOS indicates Safety Organizing Scale; n, the number of registered nurse respondents.

unit on SOS has a highly significant F-statistic (F = 2.73, $P < 0.001$).

The 2 forms of the intraclass correlation, referred to as ICC(1) and ICC(2),[38] provide omnibus indices of homogeneity and are calculated from a one-way ANOVA in which the SOS is the dependent variable and unit membership is the independent variable. ICC(1) can be interpreted as the proportion of total variance that is explained by unit membership with values ranging from −1 to +1 and values between 0.05 and 0.30 being most typical. ICC(1) is computed as

$$\text{ICC}(1) = \frac{\text{MSB} - \text{MSW}}{\text{MSB} + [(\text{k} - 1) * \text{MSW}]}$$

where MSB is between-group mean square, MSW is within-group mean square, and k is average unit size The ICC(1) value for the SOS is 0.20. Whereas ICC(1) provides an estimate of the reliability of a single RN′s assessment of the unit mean, ICC(2) provides an overall estimate of the reliability of unit means. The ICC(2) is computed as (MSB − MSW)/MSB. The closer ICC(2) is to 1.00, the more reliably nursing units can be distinguished based on individual RNs' perceptions of the SOS with values equal to or above 0.70 being acceptable. The ICC(2) value for the SOS is 0.74. In sum, the results of these 4 analyses strongly support the idea that the SOS reflects a unit-level construct and aggregation of individual responses is justified. Given this strong support, we created the unit-level SOS as the average of all the individual responses.

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

Criterion Validity

We assessed criterion validity based on hypothesized relationships between the unit-level SOS and its proposed causes and effects. First, safety organizing emerges when leaders create a supportive context that makes it safe for RNs to engage in the behaviors that comprise the SOS (eg, discussing errors).[12,13,22–24] Specifically, these behaviors should be more prevalent when RNs trust managers and administrators to treat them fairly and otherwise act with integrity.[10,22] Therefore, we expected that trust in manager would be positively related to the SOS. Second, committed RNs should be sufficiently motivated to engage in safety organizing. Third, studies have consistently linked lower patient-to-RN ratios and performance outcomes such as mortality,[39,40] failure to rescue,[40,41] complications,[41] and length of stay.[41] Fewer patients per nurse enable RNs collectively to be more vigilant in their ongoing surveillance such that they can detect and correct emerging and manifest errors in a timely manner that minimizes negative outcomes for patients.[27] As such, patients per RN should be negatively related to the SOS. Lastly, past case studies of HROs[23–25] have shown that the behaviors measured by the SOS are responsible for their nearly error-free performance. Thus, we expected that the SOS would be negatively related to reported medication errors and patient falls.

As previously mentioned, trust in manager and commitment were measured using survey items. Patient-to-RN ratio was measured using data provided by nurse managers. We also conducted our analyses using alternate measures of this ratio from survey data and the results remained substantively identical. Reported medication errors were defined as occurring whenever the right medication was not given to either the right patient, at the right time, in the right dose, or through the right route (eg, intravenous). Reported patient falls were defined as occurring any time a previously standing, sitting, or laying patient is found on the floor. Both measures were derived from incident reports filed at the time of a medication error or patient fall and were measured as the number of errors or falls reported on a unit for the 6 months after the collection of the survey data. In all our regression models, we also control for other factors that might be associated with the SOS, reported medication errors, and reported patient falls, including average RN tenure on the unit, the percentage of RNs on a unit with at least a bachelor's degree in nursing, the type of unit (a series of dummy variables for emergency department, intensive care, labor and pediatrics, surgery, and internal medicine), and hospital location (dummy variables for rural, suburban, and urban).

Criterion validity was examined through the regression analyses reported in Table 4.[42,43] The first model shows the effects of theorized antecedents on the SOS and the results confirm the expected relationships as trust in manager ($B = 0.164$, $P < 0.001$) and commitment ($B = 0.295$, $P < 0.001$)

TABLE 4. Multilevel Regression Analysis Results for Criterion Validity of the SOS

	SOS		Reported Medication Errors†		Reported Patient Falls†	
Variables	**Coefficient (*P*)**	**95% CI**	**Coefficient (*P*)**	**95% CI**	**Coefficient (*P*)**	**95% CI**
SOS	—	—	−0.678 (<0.001)	−0.907 to −0.448	−0.629 (<0.001)	−0.895 to −0.363
Trust in manager	0.164 (<0.001)	0.09 to 0.238	0.198 (<0.001)	0.102 to 0.294	0.06 (0.300)	−0.053 to 0.173
Commitment	0.295 (<0.001)	0.182 to 0.408	0.138 (0.107)	−0.03 to 0.308	0.47 (<0.001)	0.275 to 0.665
Patient-to-RN ratio	−0.056 (<0.001)	−0.084 to −0.027	0.101 (<0.001)	0.056 to 0.146	0.167 (<0.001)	0.125 to 0.208
Percent RNs with BSN	0.214 (0.108)	−0.047 to 0.474	−0.568 (<0.001)	−0.911 to −0.225	−0.100 (0.573)	−0.448 to 0.248
Average tenure on unit	−0.001 (0.985)	−0.020 to 0.019	0.044 (<0.001)	0.018 to 0.071	−0.008 (0.632)	−0.042 to 0.026
Emergency department	−0.361 (0.002)	−0.589 to −0.132	−0.643 (<0.001)	−0.952 to −0.334	−0.827 (<0.001)	−1.207 to −0.447
Intensive care unit	−0.051 (0.582)	−0.233 to 0.113	0.399 (0.003)	0.133 to 0.666	−0.861 (<0.001)	−1.204 to −0.518
Obstetrics and pediatrics	−0.058 (0.507)	−0.229 to 0.113	−0.212 (0.109)	−0.472 to 0.047	−1.072 (<0.001)	−1.459 to −0.685
Surgery	−0.006 (0.955)	−0.208 to 0.196	−0.035 (0.813)	−0.324 to 0.254	−0.959 (<0.001)	−1.444 to −0.474
Constant	3.044 (<0.001)	2.512 to 3.576	3.453 (<0.001)	2.444 to 4.462	1.893 (0.002)	0.689 to 3.097
Level-2 Model	**Variance (SE)**		**Variance (SE)**		**Variance (SE)**	
Constant	0.021 (0.033)		0.093 (0.051)		0.004 (0.027)	
Urban	0.021 (0.047)		0.887 (0.413)		0.575 (0.280)	
Suburban	0.021 (0.043)		0.318 (0.226)		0.589 (0.281)	
Log likelihood	−17.794		−508.239		−369.038	
n	118*		94‡		94‡	

All analyses were conducted in GLLAMM with hospital-level random effects. Including random effects accounts for the multilevel structure of the data (ie, units nested within hospitals).[42]

*The sample size is 118 resulting from missing patient-to-RN ratio data for 7 units. The units with missing data are not systematically different from the included units in terms of size (number of beds), scores on the SOS, or RN characteristics (age, education, or tenure).

†Multilevel Poisson regression models were estimated using GLLAMM to model reported medication errors and patient falls. We used Poisson rather than negative binomial models because hospital-level random effects sufficiently account for extra-Poisson variation (ie, overdispersion such that the variance is greater than the expectation).[43]

‡The sample size is 94 for these analyses because these units did not provide their reported medication error or patient fall data. The units with missing data are not systematically different from the included units in terms of size (number of beds), scores on the SOS, or RN characteristics (age, education, or tenure).

SOS indicates Safety Organizing Scale; CI, confidence interval, RN, registered nurse, SE, standard error; n, the number of nursing units included in the analysis.

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

are positively related to the SOS and patient-to-RN ratio is negatively related to the SOS (ie, more patients per nurse is associated with a lower score on the SOS, B = −0.06, $P < 0.001$). Models 2 and 3 show the SOS is negatively related to reported medication errors (B = −0.678, $P < 0.001$) and reported patient falls (B = −0.629, $P < 0.001$).

DISCUSSION

The 9-item SOS appears to be a precise, unidimensional measure of the self-reported behaviors underlying a safety culture. The factor structure closely resembles the initial content domains identified in the extensive case study literature on HROs and we substantively differentiated the SOS from 2 related concepts. The data also supported the SOS as a unit-level construct because aggregation of individual RN responses to the unit level was well justified. The strong relationship between the SOS and reported medication errors and patient falls provides additional support for its construct validity as well as its use. Consequently, the SOS may help differentiate between safe and unsafe units by examining the ongoing actions and interactions on the unit.

Although some studies suggest that supportive leadership[12] and a well-developed safety culture[44] should be associated with more reporting of errors, others[14] have not. We find that high levels of the SOS are associated with fewer reported medication errors and patient falls. We argue that these results coupled with additional data analyses suggest that our measure did not only reflect the tendency to report, but also provided a partial picture of patient safety on the units. If reporting more errors is associated with being a high-functioning (ie, safe) unit, we would expect reported medication errors and patient falls to be positively associated with other indicators of safety performance. We find that nurse managers' assessments of their unit's safety performance (a 2-item survey measure) are actually negatively associated with reported medication errors ($P < 0.01$) and patient falls ($P = 0.012$). That is, high numbers of reported medication errors and patient falls were associated with low ratings of quality of care by nurse managers.

To date, there have been few self-report measures developed that focus on the behaviors underlying a safety culture. To the extent that measures related to safety culture do exist, they have tended to focus on factors that create a context that promotes safety, including leaders, processes, and procedures that prioritize safety[14,18,45–47]; encourage open communication and reporting of errors[18,45–48]; and otherwise ensure that errors are handled appropriately[14,18,48] rather than how those delivering health care (eg, RNs) act to promote safety. Although some of the more comprehensive questionnaires[45,48] do possess items similar to the SOS, they only capture a portion of the behaviors documented in case studies of HROs. Given that a supportive context is essential for safety organizing, it would be fruitful to explore the relationship among measures of safety culture, the SOS, and safety outcomes.

Limitations

The findings of our study should be considered in light of its limitations. First, the research was conducted in a convenience sample of 13 Catholic hospitals that were all members of the same health system. Although this may limit the generalizability of our findings to all Catholic hospitals, other types of hospitals (eg, for-profit hospitals, academic medical centers) or health systems, the hospitals in our sample range in size from small (89 beds) to large (478 beds) and in location, including rural, suburban, and urban sites, both of which should increase confidence in the potential generalizability of the findings. In addition, the characteristics of the sample of RNs surveyed are nearly identical to the characteristics identified in a comprehensive national sample of RNs[49] in terms of age, education, and gender. However, further research is warranted to confirm that the SOS applies to a broader sample of Catholic hospitals as well as other hospital types and health systems.

Second, although our response rate was reasonable for a mailed questionnaire, it varied significantly across units and across hospitals. Subsequent analyses revealed that RN response rates at the unit and hospital levels were not correlated with reported medication errors, reported patient falls, or the SOS. Thus, when these analyses are coupled with the fact that the RNs responding to our questionnaire did not differ in any significant manner from nonrespondents, we are confident that our results are not systematically biased.

Third, our measure was validated using a sample composed exclusively of RNs. Although RNs play an important role in determining unit and hospital safety[9,17,27] and their scores on the SOS were strongly associated with fewer reported medication errors and patient falls in this study, safety is substantially determined by a much broader array of care providers (eg, doctors). Moreover, a number of impediments to patient safety (eg, diagnostic errors) are most strongly influenced by doctors. Therefore, future studies need to examine the applicability of the SOS to doctors and other care providers as well as other indicators of patient safety.

Fourth, although the SOS was linked to fewer reported medication errors and patient falls, it has been argued that reports of errors and falls are not an adequate measure of error rates.[50] Therefore, before the SOS can conclusively be determined to be associated with reductions in medical error, subsequent work should build on our strong initial findings and test criterion validity with other methods (eg, direct observation, prospective clinical surveillance) assessing explicitly defined errors and adverse events.[50]

Despite these limitations, the SOS fills an important gap in assessing the behavioral underpinnings of a unit's safety culture. The development of the SOS follows calls for investigating the behaviors facilitating error prevention[17] and the content of this measure is closely aligned with case studies documenting the association of such behaviors with high levels of safety in medical settings[23] and HROs.[21,22,24,25] The SOS items are also "actionable" in the sense that they entail concrete behaviors that can readily be changed by RNs and influenced by nurse managers through how they manage their employees on a daily basis.

In summary, the SOS makes 2 key contributions. First, it provides a self-report measure of the behaviors that lead to the emergence of a safety culture and is strongly associated with fewer reported medication errors and patient falls. Second, the

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

SOS is also a useful tool for understanding both how and under what conditions interventions designed to improve safety culture and safety outcomes such as improving patient-to-RN ratios[39–41] and safety-conscious leadership[13,14,17] are effective. That is, such interventions are likely to generate improvements by fostering safety organizing or, at least, their benefits should be enhanced in the presence of high levels of safety organizing. Future research could explicitly test the sensitivity of the SOS to safety-oriented design, leadership, or even technologic interventions and its subsequent effects on safety outcomes.

ACKNOWLEDGMENTS

The authors thank Bruce Cooil, Fred Feinberg, Jody Hoffer Gittell, Ryan Quinn, Sara Singer, Anita Tucker, and Klaus Weber for insightful comments on earlier drafts of the manuscript. The authors also thank Deputy Editor Lisa Meredith and the journal's 3 anonymous reviewers for thoughtful and constructive comments that substantially strengthened the quality of the manuscript.

REFERENCES

1. Institute of Medicine. Committee on Quality of Health Care in America. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 1999.
2. Institute of Medicine. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System of the 21st Century*. Washington, DC: National Academy Press; 2001.
3. McDonald CJ, Weiner M, Hui SL. Deaths due to medical errors are exaggerated in Institute of Medicine report. *JAMA*. 2000;284:93–94.
4. Leape LL. Institute of Medicine medical error figures are not exaggerated. *JAMA*. 2000;284:95–97.
5. Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA*. 2001;286:415–420.
6. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado in 1992. *Med Care*. 2000;38:261–271.
7. Brennan TA, Leape LL, Laird NM, et al. Incidence of adverse events and negligence care in hospitalized patients. *N Eng J Med*. 1991;324:370–376.
8. Thomas EJ, Studdert D, Newhouse JP, et al. Costs of medical injuries in Colorado and Utah in 1992. *Inquiry*. 1999;36:255–264.
9. Leape LL, Bates DW, Cullen DJ, et al. Systems analysis of adverse drug events. *JAMA*. 1995;274:35–43.
10. Reason, J. *Managing the Risks of Organizational Accidents.* Burlington, VT: Ashgate; 1997.
11. Cooper MD. Towards a model of safety culture. *Safety Sci.* 2000;36:111–136.
12. Edmondson AC. Learning from mistakes is easier said than done: group and organizational influences on the detection and correction of human error. *J Appl Behav Sci*. 1996;32:5–28.
13. Thomas EJ, Sexton JB, Neilands TB, et al. The effect of executive walk rounds on nurse safety climate attitudes: a randomized trial of clinical units. *BMC Health Serv Res*. 2005;5:28–36.
14. Naveh E, Katz-Navon T, Stern Z. Treatment errors in healthcare: a safety climate approach. *Manage Sci*. 2005;51:948–960.
15. Sexton JB, Thomas EJ, Helmreich RL, et al. *Frontline Assessments of Healthcare Culture: Safety Attitudes Questionnaire Norms and Psychometric Properties.* Austin, TX: The University of Texas Center of Excellence for Patient Safety Research and Practice; 2004. Technical Report No. 04–01. Grant No. 1PO1HS1154401. Sponsored by the Agency for Healthcare Research and Quality.
16. Sexton JB. *A Matter of Life or Death: Social Psychological and Organizational Factors Related to Patient Outcomes in the Intensive Care Unit* [dissertation]. Austin, TX: University of Texas; 2002.
17. Institute of Medicine. Committee on the Work Environment for Nurses and Patient Safety Board on Health Care Services. *Keeping Patients Safe: Transforming the Work Environment of Nurses*. Washington, DC: The National Academies Press; 2004.
18. Singer SJ, Gaba DM, Geppert JJ, et al. The culture of safety: results of an organization-wide survey in 15 California hospitals. *Qual Saf Health Care*. 2003;12:112–118.
19. Mitchell PH, Shortell SM. Adverse outcomes and variations in organization of care delivery. *Med Care.* 1997;35:NS19–NS32.
20. Sexton JB, Thomas EJ, Helmreich RL. Error, stress, and teamwork in medicine and aviation: cross sectional surveys. *BMJ*. 2000;320:745–749.
21. Weick KE, Roberts KH. Collective mind in organizations: heedful interrelating on flight decks. *Adm Sci Q*. 1993;38:357–381.
22. Schulman PR. The negotiated order of organizational reliability. *Adm Soc*. 1993;25:353–372.
23. Roberts KH, Madsen P, Desai V, et al. A case of the birth and death of a high reliability healthcare organization. *Qual Saf Health Care*. 2005; 14:216–220.
24. Weick KE, Sutcliffe KM. *Managing the Unexpected: Assuring High Performance in an Age of Complexity.* San Francisco, CA: Jossey-Bass; 2001.
25. Weick KE, Sutcliffe KM, Obstfeld D. Organizing for high reliability: processes of collective mindfulness. In: Staw BM, Cummings LL, eds. *Research in Organizational Behavior*. Greenwich, CT: JAI Press, Inc; 1999:81–123.
26. Wilson KA, Burke CS, Priest HA, et al. Promoting health care safety through training high reliability teams. *Qual Saf Health Care*. 2005;14: 303–309.
27. Aiken LH, Clarke SP, Cheung RB, et al. Educational levels of hospital nurses and surgical patient mortality. *JAMA*. 2003;290:1617–1623.
28. Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol*. 1997;50:1129–1136.
29. DeVellis RF. *Scale Development: Theory and Applications*. Thousand Oaks, CA: Sage; 2003.
30. Arbuckle JL. *AMOS 5.* Chicago, IL: SmallWaters Corp; 2003.
31. Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107:238–246.
32. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model*. 1999;6:1–55.
33. Bollen KA. *Structural Equations With Latent Variables*. New York, NY: John Wiley & Sons; 1989.
34. Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-step approach. *Psychol Bull*. 1988;103: 411–423.
35. Meyer JP, Allen NJ. *Commitment in the Workplace: Theory, Research, and Application*. Thousand Oaks, CA: Sage; 1997.
36. Klein KJ, Conn AB, Smith DB, et al. Is everyone in agreement? An exploration of within-group agreement in employee perceptions of the work environment. *J Appl Psychol*. 2001;86:3–16.
37. James LR, Demaree RG, Wolf G. Estimating within-group interrater reliability with and without response bias. *J Appl Psychol.* 1984;69:85–98.
38. Bliese PD. Within-group agreement, non-independence, and reliability: implications for data aggregation and analysis. In: Klein KJ, Kozlowski SWJ, eds. *Multilevel Theory, Research, and Methods in Organizations*. San Francisco, CA: Jossey-Bass; 2000:349–381.
39. Rothberg MB, Abraham I, Lindenauer PK, et al. Improving nurse-to-patient ratios as a cost-effective safety intervention. *Med Care*. 2005; 43:785–791.
40. Aiken LH, Clarke SP, Sloane DM, et al. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA*. 2002; 288:1987–1993.
41. Needleman J, Buerhaus P, Mattke S, et al. Nurse-staffing levels and the quality of care in hospitals. *N Engl J Med*. 2002;346:1715–1721.
42. Rabe-Hesketh S, Pickles A, Skrondal A. *GLLAMM Manual*. Technical Report 2001/01, Department of Biostatistics and Computing, Institute of Psychiatry, King's College, University of London. Available at: www.gllamm.org/.
43. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata.* College Station, TX: Stata Press; 2005.
44. Naveh E, Katz-Navon T, Stern Z. Readiness to report medical treatment errors: the effects of safety procedures, safety information, and priority of safety. *Med Care*. 2006;44:117–123.
45. Sorra JS, Nieva VF. *Hospital Survey on Patient Safety Culture*. AHRQ Publication No. 04-0041. Rockville, MD: Agency for Healthcare Research and Quality; 2004.

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

46. Weingart SN, Farbstein K, Davis RB, et al. Using a multihospital survey to examine the safety culture. *Jt Comm J Qual Saf*. 2004;30:125–132.
47. Pronovost PJ, Weast B, Holzmueller CG, et al. Evaluation of the culture of safety: survey of clinicians and managers in an academic medical center. *Qual Saf Health Care*. 2003;12:405–410.
48. Sexton JB, Helmreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res*. 2006;6:44–53.
49. *The Registered Nurse Population*. Rockville, MD: US Department of Health and Human Services; 2000.
50. Thomas EJ, Peterson LA. Measuring errors and adverse events in health care. *J Gen Intern Med*. 2003;18:61–67.

APPENDIX

APPENDIX Means, Standard Deviations, and Missing Values for the 9 SOS Items

Item*	Mean†	SD	Percent Missing	Number Missing	Alpha if Item Deleted‡	Total n
Q1 We have a good "map" of each other's talents and skills	5.05	1.20	1.10	18	0.87	1685
Q2 We talk about mistakes and ways to learn from them	4.95	1.28	1.18	20	0.86	1685
Q3 We discuss our unique skills with each other so we know who on the unit has relevant specialized skills and knowledge	4.76	1.41	0.65	11	0.86	1685
Q4 We discuss alternatives as to how to go about our normal work activities	4.29	1.41	0.53	9	0.86	1685
Q5 When giving report to an oncoming nurse, we usually discuss what to look out for	5.55	1.20	1.37	23	0.87	1685
Q6 When attempting to resolve a problem, we take advantage of the unique skills of our colleagues	5.46	1.17	0.42	7	0.86	1685
Q7 We spend time identifying activities we do not want to go wrong	4.87	1.23	1.01	17	0.86	1685
Q8 When errors happen, we discuss how we could have prevented them	5.16	1.15	0.65	11	0.86	1685
Q9 When a patient crisis occurs, we rapidly pool our collective expertise to attempt to resolve it.	5.86	1.10	1.01	17	0.87	1685

Q1 through Q9 indicates the question number associated with the scale.

*The instructions read "The following questions ask you to assess the degree to which you and the other RNs with which you currently and primarily work engage in certain behaviors and practices. By work unit, we mean your current hospital unit (eg, Cardiac Intensive Care Unit)." The stem question asks "To what extent do the following characterize your current work unit?"

†Scale: 1 = not at all, 2 = to a very limited extent, 3 = to a limited extent, 4 = to a moderate extent, 5 = to a considerable extent, 6 = to a great extent, 7 = to a very great extent.

‡These values are offered to those who might be interested in developing a short-form version of the SOS. We encourage use of the full-scale to capture the richness of all behaviors described.

SOS indicates Safety Organizing Scale; SD, standard deviation; n, the number of registered nurse respondents.

© 2006 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.