An Overview of Patient Safety Climate in the VA

*Christine W. Hartmann, Amy K. Rosen, Mark Meterko,
Priti Shokeen, Shibei Zhao, Sara Singer, Alyson Falwell,
and David M. Gaba*

Objective. To assess variation in safety climate across VA hospitals nationally.

Study Setting. Data were collected from employees at 30 VA hospitals over a 6-month period using the Patient Safety Climate in Healthcare Organizations survey.

Study Design. We sampled 100 percent of senior managers and physicians and a random 10 percent of other employees. At 10 randomly selected hospitals, we sampled an additional 100 percent of employees working in units with intrinsically higher hazards (high-hazard units [HHUs]).

Data Collection. Data were collected using an anonymous survey design.

Principal Findings. We received 4,547 responses (49 percent response rate). The percent problematic response—lower percent reflecting higher levels of patient safety climate—ranged from 12.0–23.7 percent across hospitals (mean = 17.5 percent). Differences in safety climate emerged by management level, clinician status, and workgroup. Supervisors and front-line staff reported lower levels of safety climate than senior managers; clinician responses reflected lower levels of safety climate than those of nonclinicians; and responses of employees in HHUs reflected lower levels of safety climate than those of workers in other areas.

Conclusions. This is the first systematic study of patient safety climate in VA hospitals. Findings indicate an overall positive safety climate across the VA, but there is room for improvement.

Key Words. Patient safety, safety climate, high-reliability organization

The health care industry attempts to cure patients while avoiding problems and negative outcomes resulting from the processes of care. In many ways, health care strives to be a "high-reliability" enterprise. High-reliability organizations (HROs), which include the aviation and nuclear power industries, successfully perform highly complex and repetitive tasks while avoiding disastrous events. The reliability of HROs is attributed in part to having a strong safety climate, and several different formulations of key principles of

HROs have been introduced into health care (Gaba 2001; Weick and Sutcliffe 2001, 2003; Roberts and Tadmor 2002; Rall and Dieckmann 2005; Gaba, Singer, and Rosen 2006). Research has shown several components to be important in achieving a high-safety climate: (1) attitudes and perceptions must be uniformly safety oriented (Hofmann and Morgeson 1999; Sexton, Thomas, and Helmreich 2000b; Katz-Navon, Naveh, and Stern 2005); (2) procedures, structures, and resources supportive of safety are in place (Shekelle et al. 2000; Sexton, Thomas, and Helmreich 2000b; Zohar 2002; Tucker and Edmondson 2003); (3) there is adequate safety training for personnel (Roberts 1990; Morey et al. 2002); and (4) auditing of clinical processes and safety standards takes place (Welsh, Pedot, and Anderson 1996; Weingart, Ship, and Aronson 2000).

As the nation's largest integrated health system, the VA provides unique opportunities to study patient-safety climate. With the current emphasis on moving from a punitive climate within hospitals toward a system-wide climate of safety within health care in order to reduce medical errors (Bagian and Perlin 2000; Reason 2000; IOM 2001; Kizer 2001), assessing the status of the VA's safety climate is essential. However, little is known about the safety climate of the VA because previous surveys in the VA on this topic suffer from a number of methodological limitations, including small sample sizes and nonrandomized samples. Results have not been published in the peer-reviewed literature (Landesman and McKnight 2006). This study represents the first systematic investigation of safety climate within VA hospitals using HRO theory as the theoretical base and allows for a more comprehensive understanding of factors affecting patient safety than has been possible to date.

Address correspondence to Christine W. Hartmann, Ph.D., Center for Health Quality, Outcomes and Economic Research, A VA Center of Excellence, Bedford, MA and the Department of Health Policy and Management, School of Public Health, Boston University, Boston, MA, Center for Health Quality, Outcomes and Economic Research, Bedford VAMC (152), 200 Springs Road, Bedford MA 01730. E-mail: christine.hartmann@va.gov. Amy Rosen, Ph.D., is with the Center for Health Quality, Outcomes and Economic Research, A VA Center of Excellence, Bedford, MA, and the Department of Health Policy and Management, School of Public Health, Boston University, Boston, MA. Mark Meterko, Ph.D., is with the Center for Organization, Leadership and Management Research, A VA Center of Excellence, Boston, MA. Priti Shokeen, M.S., and Shibei Zhao, M.P.H., are with the Center for Health Quality, Outcomes and Economic Research, A VA Center of Excellence, Bedford, MA. Sara Singer, Ph.D., is with the Department of Health Policy and Management, Harvard School of Public Health. Alyson Falwell, M.P.H., and David M. Gaba, M.D., are with the Center for Health Policy and Center for Primary Care and Outcomes Research (CHP/PCOR), Stanford University, Stanford, CA.

We used a previously validated instrument (Singer et al. 2007) to measure patient safety climate in VA hospitals. Using HRO theory as a conceptual framework, the principal objective of our study was to assess the levels of safety climate in VA hospitals. We hypothesized the following:

(1) there will be considerable variability in levels of safety climate within hospitals, among job types and workgroups (i.e., units or departments), as well as across VA hospitals;
(2) front-line workers and clinicians, because of their direct exposure to medical errors, will perceive lower levels of safety climate than managers or nonclinicians; and
(3) personnel working in units facing intrinsically higher hazards ("high-hazard units" [HHUs], where the work, interventions, and drugs used have high intrinsic potential lethality) will perceive lower levels of safety climate than personnel in other units.

METHODS

We used the Patient Safety Climate in Healthcare Organizations (PSCHO) survey to measure the safety climate of participating hospitals. The PSCHO was developed by the Patient Safety Culture Institute (PSCI) at the VA Palo Alto Health Care System and Stanford University (Singer 2007). Evaluation has demonstrated generally favorable psychometric properties (Colla et al. 2005; Singer et al. 2007b). The PSCHO contained a total of 42 items related to safety climate and six demographic questions. Responses to all items were close-ended. Each safety climate item used a 5-point, neutral mid-point Likert scale ("strongly disagree" = 1 to "strongly agree" = 5) with an additional "not applicable" option.

Because the literature on HROs emphasizes the necessity of a "preoccupation with failure" and the importance of a highly uniform safety climate, the PSCHO instrument is scored to reflect the percentage of "*problematic* responses"—answers that suggest an *absence* of safety climate. A lower percentage of problematic responses indicates a higher level of safety climate. Identifying frequencies of problematic response focuses attention on areas of safety climate that may be susceptible to error and may present opportunities for improvement. Weighting each survey item equally, we calculated overall and item-specific frequencies of problematic response by hospital, job category, and workgroup. In addition, because a neutral mid-point response

(i.e., neither agree nor disagree) to an item could potentially be characterized as problematic with respect to safety climate (Singer et al. 2003), we calculated the frequency of neutral responses to all items and scales and for scales by management category.

Before conducting this study, approval was granted from the relevant Institutional Review Boards for all participating hospitals.

Survey Sample

To minimize selection bias, we conducted a complex recruitment process to accrue participating hospitals. All 117 VA acute care hospitals were divided into four performance strata based on the Agency for Healthcare Research and Quality (AHRQ)'s Patient Safety Indicators (PSIs) (Shojania et al. 2001; Romano et al. 2003): low, medium, high, and other. To do this, we calculated two PSI rates for each hospital (one for each of two categories of PSIs created through a factor analysis: complications related and medical/surgical) and ranked hospitals for each group of PSIs separately, dividing each ranking into quartiles. We assigned hospitals to strata based on the quartiles, and those whose ranking in the two PSI groups differed by more than one quartile were assigned to the "other" category. Within each stratum, hospitals were recruited in random order. We recruited a total sample of 30 hospitals for participation in the study—eight from the high, medium, and other categories; six from the low category. Details of the hospital recruitment strategy have been described elsewhere (Rosen et al. in press).

Data from the 2004 American Hospital Association Annual Survey of Hospitals provided information on hospitals' teaching status, bedsize, and geographic region, and the 2005 Bureau of Health Professions' Area Resource File provided information about metropolitan location. To calculate hospitals' case mix, we used AHRQ's comorbidity software on 2005 VA discharge-level data (Patient Treatment File) (AHRQ 2003). These data were aggregated to the hospital level to calculate a hospital's case mix. Data from the VA's 2005 Decision Support System files were used to calculate nurse staffing ratios. Almost all hospitals in the sample were located in metropolitan areas (90 percent, $n = 26$), and over half (53 percent, $n = 16$) were major teaching hospitals (members of the Council of Teaching Hospitals of the Association of American Medical Colleges). Fifty-five percent ($n = 16$) were large (bedsize > 250), 21 percent ($n = 6$) were medium (bedsize $= 100–249$), and 24 percent ($n = 7$) were small (bedsize ≤ 99). At least one hospital from each of the four U.S. geographic regions (east, south, midwest, and west) was represented

in each of the size categories, except for small, which was missing a hospital from the "west."

In all participating hospitals, we sampled 100 percent of the senior managers (department head and above), 100 percent of hospital-based physicians, and a random 10 percent sample of other employees. In addition, at 10 randomly selected hospitals, we sampled 100 percent of employees from HHUs: operating room (OR) (including postanesthesia care unit [PACU]), intensive care unit (ICU), emergency department (ED), and intravenous chemotherapy administration unit. We oversampled senior managers because of their small number and physicians because of their typically low response rates (Berk 1985). Because nurses comprise such a large fraction of hospital employees, it was not necessary to oversample them. Our original sampling frame consisted of 10,837 employees' names. Out of these, 907 were removed because the individual was no longer working at the hospital or had died. An additional 621 individuals were eliminated because they declined to participate using a response postcard, leaving a final sampling frame of 9,309 individuals.

Administration of Survey

Survey printing, distribution, and collection were conducted by an outside vendor specializing in survey administration. Surveys were delivered to each hospital via U.S. mail and distributed internally by each site's project coordinator. Between December 2005 and May 2006, participants received up to three surveys in waves spaced approximately 5 weeks apart. Participation was voluntary and all responses were anonymous. Each survey packet contained a cover letter with instructions, the survey instrument, a business reply return envelope, and, for waves 1 and 2, a survey completion postcard, which allowed for anonymous tracking of responses and declines for subsequent mailings.

Weighting of Data

We generated combined weights for each workgroup within each hospital in order to reflect the original sampling frame accurately. Combined weights were calculated by multiplying the sampling weight by the nonresponse weight. All analyses were conducted using the combined weights.

Sampling Weight. A sampling weight was calculated for each workgroup *within each hospital* in order to account for variation in size from the original

sampling target. For each of the 100 percent samples (senior administrators, physicians, and employees of HHUs), the sampling weight was set as 1.0. For the other employees, where the target sample size was 10 percent, the sampling weight was close to 10 (the numerator consisted of the total number of employees in this category as reported by each hospital, while the denominator contained the number of employees in our original sampling frame).

NonResponse Weight. To account for varying response rates within workgroups, we assigned a nonresponse weight for each workgroup *within each hospital*, calculated using the inverse of the response rate. Response rates were calculated using the number of responses received for a specific hospital's workgroup over the number of employees in this workgroup (based on our final, deliverable sample).

Analysis of Data

Psychometric Analysis. To assess the psychometric properties of the revised PSCHO instrument, respondents were randomly divided into a derivation sample ($n = 2,252$) and a validation sample ($n = 2,252$). We applied multitrait analysis (MTA) to the data in the derivation sample to assess both the reliability and validity of the multi-item scales (Campbell and Fiske 1959), beginning with item-to-scale assignments based on our conceptual model of the key elements of safety climate. Item content and scale interpretability were considered, and reliability and validity statistics were examined. Items were dropped from a given scale to improve scale reliability or decrease interscale correlations. MTA was applied to the revised scale structure, and the cycle was repeated until no further improvements in scale psychometrics could be achieved. We found empirical support for 11 of the conceptually based safety climate scales, which we labeled "senior leadership," "resources for safety," "facility characteristics," "workgroup leadership," "workgroup norms," "workgroup recognition," "learning," "fear of blame," "psychological safety," "problem responsiveness," and "outcomes." These align generally with the components necessary to achieve a strong safety climate outlined above. For example, "resources for safety" items reflect the need for supportive structures and procedures, while "learning" items reflect the need for auditing of procedures.

To test the stability and robustness of the MTA results, we applied confirmatory factor analysis (CFA) (Bollen 1989; Bryant and Yarnold 1995) in the validation sample using the final scale structure produced by the MTA.

Survey Data Analysis. Our primary outcome was percent of problematic response to survey items. We conducted a correlation analysis and, to further investigate the effects of management level on problematic response frequencies, constructed a multiple regression model treating each hospital as a primary sampling unit. These results are reported where they provide additional information or support data presented in the tables.

To assess the potential limitation of differences in problematic response by time to response, we compared differences in mean problematic response by wave using 95 percent confidence intervals and found no significant differences.

All analyses were performed at the individual level using $SAS^{©}$, version 9.1 (SAS Institute, Cary, NC).

RESULTS

Survey Response

One hospital returned data for only the physician sample and no other employees and was dropped from analyses. All results reported reflect data for 29 hospitals. A total of 4,547 surveys was returned, for an overall response rate of 49 percent. Job category was related to response rate, with senior managers having the highest (68 percent) and physicians the lowest (37 percent); excluding physicians, the response rate was 57 percent. The response rate also varied by hospital, ranging from 26 to 73 percent.

Psychometric Properties of PSCHO

Eighty-eight percent of the items had correlations of 0.40 or higher with their respective scale scores (adjusted for overlap) from the 11 scales resulting from the MTA, suggesting adequate item internal consistency (Kerlinger 1973; Ware et al. 1997). Further, in 329 out of 400 comparisons (82 percent), the correlations between items and their hypothesized scales were significantly higher than their correlations with any other scale; they were higher, though not significantly so, in an additional 37 comparisons. Altogether, appropriate discriminant validity was observed for 92 percent of item-to-scale correlations. Internal consistency reliabilities ranged from 0.61 to 0.89.

We did not find strong support for the hypothesized "fear of shame" scale, which consists of three items related to individual feelings of shame associated with mistakes and help seeking that could affect willingness to come

forward with safety concerns in a timely manner. Internal consistency reliability for this scale was low (0.46), and none of the items achieved item-to-respective scale correlations above the 0.40 criterion. We did not treat these items as a scale in the analyses.

Results of the CFA in the validation sample supported the scale structure suggested by the MTA in the derivation sample. The root mean square error of approximation was 0.065, and Bentler's normed comparative fit index was 0.98. Both values were in the range indicative of a good fit of the model to the data (Bentler and Bonett 1980; Hu and Bentler 1999).

Demographic Information

Respondents were evenly divided between males and females (Table 1). Most, 68 percent, were between the ages of 41 and 60. Thirteen percent of respondents were senior managers. Over one-third (36 percent) were physicians. Finally, 17 percent were HHU employees.

Problematic Response by Item and Scale

Results for individual items grouped by scale are displayed in Table 2. The mean percent problematic response across all hospitals was 17.5, ranging from 12.0 to 23.7. Across all hospitals, individual item percent problematic response ranged from 4.4 to 49.6, with a mean of 18.1. The mean problematic response across all scales was 18.7 percent. Thirty-four items had a percent problematic response of 10 percent or greater, and 10 had a problematic response of 25 percent or greater. The scale with the highest α-coefficient was "senior leadership" (.89).

Overall, 50 percent of respondents indicated that their workgroups did not recognize individual safety achievement through rewards, nor were they rewarded for timely action to identify a serious mistake (41 percent). Forty-three percent reported never to have witnessed a coworker perform an unsafe act during patient care. Seven percent believed that asking for help was a sign of incompetence, and 4 percent indicated that they would not report a mistake with significant consequences that no one had noticed.

Across all items, an average of 19.9 percent of respondents selected the neutral mid-point of the scale. The percent of neutral responses ranged from 5.2 ("If I make a mistake that has significant consequences and nobody notices, I do not tell anyone about it") to 39.3 percent ("Clinicians who make serious mistakes are usually punished"). This latter item is part of the "fear of blame" scale, designed to capture a punitive climate. The average percent

Table 1:   Demographic Information of Respondents*

	Overall (%)	Women (%)	Men (%)
Overall	100.0	50.4	49.6
Job category			
Senior managers	12.8	3.8	8.9
Supervisor	20.2	8.2	11.9
Other employees	67.0	38.2	28.8
Age			
<30	4.1	3.1	1.1
31–40	15.8	8.8	7.0
41–50	31.0	17.6	13.4
51–60	36.6	16.9	19.8
>60	12.4	4.2	8.3
Clinician status			
Physician[†]	36.2	9.8	26.5
Nurse[‡]	22.8	19.3	3.5
Other clinician[§]	6.8	3.7	3.1
Nonclinician[¶]	34.2	17.7	16.5
HHU[‖]			
HHU	17.2	9.9	7.2
Other	82.9	40.3	42.6

*Totals may not sum to 100 due to rounding.

[†]Physician = physician, resident, intern, or fellow.

[‡]Nurse = RN, RNP, LVN, and LPN.

[§]Other clinician = nursing assistant, pharmacist, physician assistant, physical therapist, respiratory therapist, audiologist, psychologist, speech pathologist, or occupational therapist.

[¶]Non-clinician = food, laboratory and radiology technician, ward clerk, patient escort, housekeeping, or other staff categories.

[‖]HHU = operating room (including postanesthesia care unit), intensive care unit, emergency department, and intravenous chemotherapy administration unit.

HHU, high-hazard units.

problematic response plus neutral response over all items was 37.9 percent. Items rated by participants as "not applicable" were not included in our analyses.

Scale Problematic Response by Position and Job Type

We examined the mean percent problematic response by scale for senior managers compared with nonsenior managers (i.e., all others) as well as clinicians versus nonclinicians (Table 3). There was a statistically significant difference between the mean problematic response for senior managers (9.8 percent) and that for all others (18.3 percent). Senior managers had

Table 2:   Responses* by Scale

Scales and Text of Item (Cronbach's α-Coefficient)	% Problematic	% Problematic + % Neutral
Senior leadership (0.89)	18.4	38.5
Good communication flow exists up and down the chain of command regarding patient safety issues	17.2	29.3
Senior management supports a climate that promotes patient safety	12.2	25.1
Senior management has a clear picture of the risks associated with patient care	17	34.6
Senior management has a good idea of the kinds of mistakes that actually occur in this facility	21.1	44.2
Senior management considers patient safety when program changes are discussed	14.4	40.9
Patient safety decisions are made by the most qualified people regardless of rank or hierarchy	28.8	56.8
Resources for safety (0.76)	16.8	34.4
I am provided with adequate resources (personnel, budget, and equipment) to provide safe patient care	25	38.7
I have enough time to complete patient care tasks safely	21	36.6
I have received sufficient training to enable me to address patient safety problems	8.5	22.7
This facility devotes sufficient resources to follow-up on identified safety problems	12.6	39.6
Facility characteristics (0.68)	10.5	36.5
Compared with other facilities in the area, this facility cares more about the quality of patient care it provides	12.5	42.9
Overall the level of patient safety at this facility is improving	8.5	30.0
Workgroup leadership (0.68)	22.3	44.8
Management in my unit helps me overcome problems that make it hard for me to provide safe patient care	25.4	51.1
In my unit, management puts safety at the same level of importance as meeting the schedule and productivity	20.3	42.2
Whenever pressure builds up, management in my unit wants us to work faster, even if it means taking shortcuts that might negatively affect patient safety	21.1	41.1
Workgroup norms (0.72)	11.8	28.5
My unit takes the time to identify and assess risks to ensure patient safety	8.7	19.3
My unit does a good job managing risks to ensure patient safety	7.4	18.9
I have learned how to do my own job better by learning about mistakes made by my coworkers	14.1	36.4
In my unit, there is significant peer pressure to discourage unsafe patient care	21.5	44.9
In my unit, anyone found to intentionally violate standards or safety rules is corrected	10.7	28.9
Deliberate violations of standard operating procedures are rare in my unit	8.6	22.5

Table 2. *Continued*

Scales and Text of Item (Cronbach's α-Coefficient)	% Problematic	% Problematic + % Neutral
Workgroup recognition (0.71)	38.6	62.4
I am rewarded for taking quick action to identify a serious mistake	40.6	71.8
My unit recognizes individual safety achievement through rewards	49.6	72.3
My unit provides training on teamwork in order to improve patient care performance and safety	25.6	43.2
Fear of shame[†] *(0.46)*	13.7	22.6
Asking for help is a sign of incompetence	6.6	11.8
If I make a mistake that has significant consequences and nobody notices, I do not tell anyone about it	4.4	9.4
Telling others about my mistakes is embarrassing	30.2	46.6
Learning (0.66)	10.9	34.8
Mistakes have led to positive changes in my unit	12.4	41.5
My performance is evaluated against defined safety standards	12.2	38.1
In my unit, patient safety problems and errors are communicated to the right people so that the problem can be corrected	8.2	24.8
Fear of blame (0.61)	23.9	57.5
If people find out that I made a mistake, I will be disciplined	27.2	55.1
Clinicians who make serious mistakes are usually punished	20.7	59.9
Psychological safety (0.75)	14.7	30.7
Staff feel comfortable questioning the actions of those with more authority when patient safety is at risk	22.3	43.2
Staff freely speak up if they see something that may negatively affect patient care	12	26.7
I am comfortable reporting safety concerns without fear of being punished by management	10	22.2
Problem responsiveness (0.84)	13.9	37.7
Bringing patient safety concerns to management's attention usually results in the problem being addressed	18.3	39.0
In my unit, we identify and fix safety problems before an incident actually occurs	12.7	35.5
When I take time to communicate about patient safety problems there is appropriate follow-up	14.1	38.7
Our process of accident incident investigation is effective at identifying root causes	10.7	37.5
Outcomes (0.63)	28.5	42.3
In the last year, I have witnessed a coworker do something that appeared to me to be unsafe for the patient	31.1	46.3
I have never witnessed a coworker do something that appeared to me to be unsafe patient care	43	58.6
In the last year, I have done something that was not safe for the patient	11.3	22.1
Overall average	18.1	37.9

*Responses weighted for sampling and for nonresponse.

[†]Hypothesized but not supported as a scale by psychometric data.

Table 3:   Scale Responses by Management Level, Clinician Status, and Workgroup

Scale	Responses† by Management Level & Clinician Status (% Problematic)				Responses† by Workgroup (% Problematic)						
	Non-senior Managers	Senior Managers	Non-clinicians	Clinicians	Ambulatory	ED+Urgent	ICU	Non-clinical	OR+PACU	Pharm+Lab	Ward
Senior leadership	18.7	7.2*	15.5	21.1	17.3	25.9	27.1	15.1	23.8	18.5	20.3
Resources for safety	16.9	11.5	14.3	18.7	15.3	22.3	21.0	14.1	17.8	15.7	20.3
Facility characteristics	10.7	2.5	8.4	12.4	9.3	16.2	13.8	6.9	15.5	14.2	12.7
Workgroup leadership	22.5	9.5*	20.5	23.7	20.5	27.9	30.4	20.7	20.6	23.8	23.5
Workgroup norms	12.0	4.3*	12.8	11.0	8.9	13.3	12.5	12.3	13.3	14.5	12.3
Workgroup recognition	39.0	16.9*	37.2	39.8	37.1	47.1	46.8	35.8	43.4	44.9	39.3
Learning	11.0	5.8	11.6	10.4	9.6§	18.7§	14.7§	11.3§	9.7§	10.1§	10.9§
Fear of blame	24.2	13.3	23.5	24.3	23.5	23.1	22.6	22.3	26.0	23.0	26.0
Psychological safety	14.9	5.8	15.7	13.9	11.8	16.4	12.9	14.5	13.7	21.4	16.0
Problem responsiveness	14.2	4.6*	12.8	15.0	11.8§	20.9§	18.4§	10.8§	17.6§	16.6§	15.6§
Outcomes	28.5	30.1	22.1	33.8	27.9	38.6	35.4	20.7	35.6	26.9	34.3
Overall average	18.3	9.8*	16.8	19.3	17.3	23.4	22.7	16.5	20.7	20.3	20.4

*Significantly different from nonsenior managers ($p<.05$, t-test).

†Responses weighted for sampling and for nonresponse.

§Significant differences found among workgroups for the scale ($p<.05$, ANOVA).

significantly lower overall means than nonsenior managers ($p < .05$) on the "senior leadership," "workgroup leadership," "workgroup norms," "workgroup recognition," and "problem responsiveness" scales. In contrast, clinicians and nonclinicians had relatively similar responses, with no statistically significant differences between their overall means. When we looked at the data by workgroup, we found statistically significant differences among workgroups for the "learning" and "problem responsiveness" scales.

We found a significant correlation between senior managers and clinicians. We constructed a multiple regression model using percent problematic response as the dependent variable and management level as the independent variable, controlling for facility- and individual-level characteristics (geographic region, teaching status, metropolitan location, bed size, nurse-to-patient ratio, case mix, gender, age, HHU, and nurse). We found that senior managers have a significantly lower percent problematic response than other job types (data not shown).

To assess the relationship between percent problematic and percent neutral responses, we investigated the percent neutral response over all items by management level by creating a dummy variable for senior managers. For all but three items, senior managers had a lower percent neutral response than other staff (data not shown). The three items for which senior managers had higher frequencies of neutral response were "My unit recognizes individual safety achievement through rewards and incentives" (28.2 percent versus 20.6 percent), "Telling others about my mistakes is embarrassing" (21.5 percent versus 15.9 percent), and "I am rewarded for taking quick action to identify a serious mistake" (34.6 percent versus 28.3 percent).

Scale Problematic Response by Workgroup

To investigate the effect of hospital workgroup on problematic response, we created seven workgroup categories based on similarity of clinical functions: ambulatory (home care, off-campus ambulatory care, and main-campus ambulatory care); ED and urgent care; ICU; nonclinical; OR and PACU; pharmacy and lab; and ward (Table 3). The three workgroups with the highest overall mean percent problematic responses (ED/urgent, mean = 23.4 percent; ICU, mean = 22.7 percent; and OR/PACU, mean = 20.7 percent) substantially overlapped with the areas categorized as HHUs in our sampling strategy. For seven of the 12 scales, the nonclinical workgroup had the lowest mean percent problematic response. Either ED/urgent or ICU had the highest percent problematic response for all scales, with the exception of the "fear of blame," "psychological safety," and "workgroup norms" scales.

Effects of Management Level within Job Type

We explored the variation in percent problematic response by management level and job type by creating four job categories: physician (including resident), nurse, other clinician, and nonclinician. Management level was broken down into three groups: senior manager, supervisor, and front-line worker. Nonclinicians had the largest percent of senior managers; 16.4 percent of responding nonclinicians were senior managers, compared with 14.7 percent of physicians, 8.9 percent of other clinicians, and 5.2 percent of nurses. Of the 240 nonclinician senior managers, 236 identified themselves as "other" (i.e., having administrative duties) on the demographic questions.

We calculated a mean percent problematic response for each group (Figure 1). Across all job categories, front-line workers had the highest mean and senior managers the lowest. Nurses had the largest variation in mean percent problematic response by management level (senior manager nurses, 9.5 percent; front-line nurses, 21.8 percent; standard deviation, 12.2 percent), closely followed by nonclinicians (senior manager nonclinicians, 7.7 percent; front-line nonclinicians, 18.1 percent; standard deviation,

Figure 1:   Variation in Problematic Response by Management Level and Job Type.
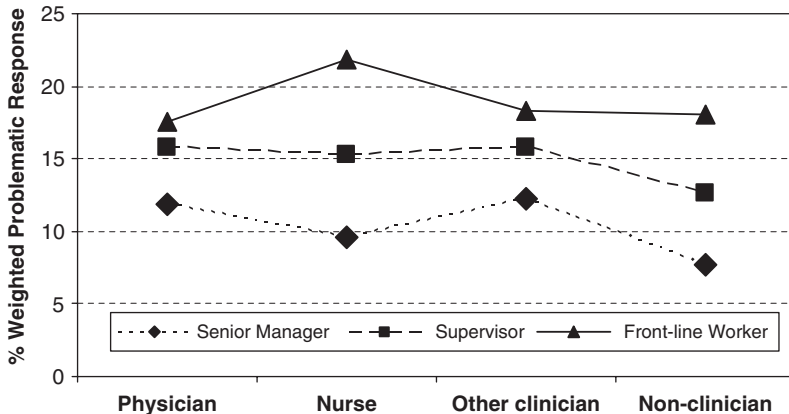
Physician = physician, resident, intern, or fellow.
RN = RN, RNP, LVN, or LPN.
Other clinician = nursing assistant, pharmacist, physician assistant, physical therapist, respiratory therapist, audiologist, psychologist, speech pathologist, or occupational therapist.
Nonclinician = food technician, lab technician, ward clerk, patient escort, radiology technician, housekeeping staff, or other.

8.1 percent). Physicians as a group and front-line workers overall had almost identical variations.

DISCUSSION

The psychometric results of our study provide strong support for 11 scales representing dimensions of safety climate. We found that VA hospitals possess relatively high levels of safety climate (83 percent of responses overall were *not* problematic), and items with the highest percent problematic response were those related to rewards and recognition of safe behavior, indicating a need for programs aimed at recognizing and rewarding contributions to improving safety. Despite a relatively high level of safety climate in general, results suggest important opportunities for improvement in VA hospitals.

A key question is the meaning of these results as defined by HRO theory. A fundamental feature of HROs is their achievement of a highly uniform safety climate. It is not enough that a majority of employees believe safety climate to be important, because the presence of even a relatively small minority of individuals who do not can raise risks within an organization that may, over time, undermine the performance of others. No minimum threshold for unsafe attitudes or practices exists that enables the differentiation of an authentic HRO from one that is not, but some notable HRO theorists have suggested that having more than 10 percent of a workforce answer items in ways that are problematic regarding safety climate is of concern (A.P. Ciavarelli, Naval Postgraduate School, personal communication; K.H. Roberts, Haas School of Business, University of California, Berkeley, personal communication). In fact, several survey studies of commercial and naval aviators have shown that in these HRO settings high uniformity of answers in favor of safety is achieved (Sexton, Thomas, and Helmreich 2000b; Gaba et al. 2003). Given these results, the level of safety climate within the VA overall has still not achieved highly reliable levels and could benefit from further study and targeted action on rewards and safe behaviors.

Comparing our results with studies outside the VA that have used the PSCHO instrument, we found similar levels of problematic response overall. The overall mean problematic response of the means across hospitals in a recent study using the PSCHO instrument conducted across 92 non-VA hospitals was 17.6 percent (Singer 2007), compared with our mean across VA hospitals of 17.7 percent. Direct item-to-item comparison is not possible here because different versions of the PSCHO instrument were used. Noteworthy

also is the increase in problematic responses when neutral is classified as problematic. (The non-VA study above had similar results, with a problematic plus neutral percent problematic response of 34.2 percent [personal communication to C.W. Hartman from S.J. Singer, Center for Health Policy/Center for Primary Care and Outcomes Research, Stanford University].) The range of 34.1 percentage points for neutral responses was over 11 percentage points lower than the range observed for the problematic response categories, suggesting that neutral was not used as a catch-all or a choice for unenthusiastic respondents. Instead, it may be important as a signal of potentially problematic areas and is worthy of further analysis.

One important finding was the difference in problematic responses between senior managers and other staff. On five of the 11 scales, workers who are not senior managers, including front-line staff who are in direct contact with patients and therefore have the most exposure to potential safety issues, had significantly higher levels of problematic response than senior managers, whose work is by nature removed from the direct patient care environment. This supports our hypothesis and indicates a possible difference in understanding of the issues affecting patient care between those providing the care and hospital administration; this interpretation is substantiated by our findings regarding management level and job type. Across all job types, senior managers had the lowest percent problematic response, with the largest differences in problematic response by management level being for nurses and the lowest for physicians, potentially because nurse managers, unlike physician managers, are less likely to have direct patient contact (Singer 2007).

The workgroup has been (Shortell et al. 1991; Shortell et al. 1994; Sexton, Thomas, and Helmreich 2000a) and is being (Pronovost et al. 2003; Pronovost et al. 2004) studied as a locus of safety climate. Specific differences in perceptions of safety climate emerged by workgroup in our study. Workers in the ICU, ED, and urgent care had higher levels of overall problematic response and higher problematic responses on the "learning" and "problem responsiveness" scales than workers in other areas, which is not surprising, given the complex and urgent nature of the tasks performed in these areas. It makes sense that areas of high complexity and hazard will yield more challenges to safety, resulting in a greater incidence of problematic response (Sexton, Makary et al. 2006). On the other hand, such areas have the closest parallels to work in other high-hazard industries in which a more uniform safety climate has been successfully achieved, suggesting that lower levels of problematic response are possible. However, the hospital as an institution is an equally important locus and represents an aggregate possessor of the inter-

acting sub-climates of its workgroups. All workgroups share the same hospital senior management, who, through the resources they control and their decision-making authority, strongly influence the safety climate that individuals and workgroups express. Therefore, it is important to investigate variation both within workgroups and overall as key indicators of overall safety climate (Gaba, Singer, and Rosen 2007).

Limitations

Although we achieved an overall response rate of almost 50 percent and found no difference in problematic response relative to the time of response, the possibility of nonresponse bias in our results cannot be dismissed. Ours was a voluntary survey administered without incentives to busy hospital staff. Although our response rate is analogous to that achieved in similar studies (Singer et al. 2003, 2007), we acknowledge that other safety climate surveys have achieved higher response rates (Sexton, Helmreich et al. 2006; Sorra, Famolaro, and Dyer 2007). Nonetheless, we used a clearly outlined research methodology to achieve this rate, and we attempted to compensate for our lowest response rate category (i.e., physicians) by sampling 100 percent of physicians. For comparisons among job-types, we adjusted for nonresponse by weighting results to account for the differential response rate of the different target populations.

We made a great effort to recruit hospitals across a broad spectrum of performance and geographic criteria (Rosen et al. in press) but our results may not be completely representative of the larger VA population. Analysis of our recruitment data shows that there was a significantly higher percentage of hospitals with low PSI rates (i.e., "high performers" in terms of patient safety) in our sample than in the nonrecruited population ($p < .05$), indicating a possible hospital selection bias. However, our efforts to recruit across the spectrum of performance improved the sample's distribution. Further, if our sample included an overabundance of high-performing hospitals (presumably those with higher levels of safety climate), then our data may actually represent a lower bound on the prevalence of a problematic safety climate within VA hospitals.

Finally, as with all cross-sectional studies, the data presented here represent only one point in time, which may be of particular concern given ongoing efforts to improve patient safety climate in VA hospitals. In addition, there may have been other factors affecting levels of problematic response that were not captured by our instrument. However, other studies using similar

instruments outside the VA have found comparable percentages of problematic response (Singer et al. 2003, 2007).

CONCLUSIONS

It is important to understand the level of safety climate in hospitals to inform improvement efforts. This study presents the first reliable and valid information about a representative sample of VA hospitals. Our findings suggest that there is room for improvement and that institutions vary in their perceptions of safety climate, and in their strengths and weaknesses. Data from VA hospitals, like those from non-VA hospitals, suggest a complex picture of safety climate and attitudes linked to a number of factors. All hospitals need to attend to differences by management level and workgroup; more information about the causes of variation is needed. Our results point to a generally positive climate of safety across all VA institutions, workgroups, and types of workers. However, for the VA, similar to other health care organizations, it still remains a challenge to devise and implement strategies and interventions that successfully create and sustain the kind of highly intense climate of safety that is thought to be a hallmark of high reliability in other domains.

ACKNOWLEDGMENTS

REFERENCES

AHRQ. 2003. "Patient Safety Indicators SAS Software Documentation." Rockville, MD: Agency for Healthcare Research and Quality (AHRQ).

Bagian, J. P., and J. B. Perlin. 2000. "Statement before the House Veterans' Affairs Subcommittee on Oversight & Investigation" [accessed on February 4, 2008]. Available at http://veterans.house.gov/hearings/schedule106/july00/7-27-00/jperlin.htm

Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88: 588–606.

Berk, M. L. 1985. "Interviewing Physicians: The Effect of Improved Response Rate." *American Journal of Public Health* 75 (11): 1338–40.

Bollen, K. A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.

Bryant, F. B., and P. R. Yarnold. 1995. "Principal-Components Analysis and Exploratory and Confirmatory Factor Analysis." In *Reading and Understanding Multivariate Statistics*, edited by L. G. Grimm and P. R. Yarnold, pp. 99–136. Washington, DC: American Psychological Association.

Campbell, D. T., and D. W. Fiske. 1959. "Convergent and Discriminant Validity by the Multitrait–Multimethod Matrix." *Psychological Bulletin* 56: 81–105.

Colla, J. B., A. C. Bracken, L. M. Kinney, and W. B. Weeks. 2005. "Measuring Patient Safety Climate: A Review of Surveys." *Quality and Safety in Health Care* 14: 364–6.

Gaba, D. M. 2001. "Structural and Organizational Issues in Patient Safety: A Comparison of Health Care to Other High-Hazard Industries." *California Management Review* 43: 83–102.

Gaba, D. M., S. J. Singer, and A. K. Rosen. 2007. "Safety Culture: Is the 'Unit' the Right 'Unit of Analysis'?" *Critical Care Medicine* 35 (1): 314–6.

Gaba, D. M., S. J. Singer, A. D. Sinaiko, J. D. Bowen, and A. P. Ciavarelli. 2003. "Differences in Safety Climate between Hospital Personnel and Naval Aviators." *Human Factors* 45 (2): 173–85.

Hofmann, D., and F. Morgeson. 1999. "Safety-Related Behavior as a Social Exchange: The Role of Perceived Organizational Support and Leader-Member Exchange." *Journal of Applied Pyschology* 84 (2): 286–96.

Hu, L., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling* 6 (1): 1–55.

IOM. 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: Institute of Medicine, National Academy Press.

Katz-Navon, T., E. Naveh, and Z. Stern. 2005. "Safety Climate in Health Care Organizations: A Multidimensional Approach." *Academy of Management Journal* 48 (6): 1075–89.

Kerlinger, F. N. 1973. *Foundations of Behavioral Research*. New York: Holt, Rinehart & Winston Inc.

Kizer, K. W. 2001. "Patient Safety. A Call to Action: A Consensus Statement from the National Quality Forum." *Medscape General Medicine* 3 (2): 10.

Landesman, A., and S. McKnight. 2006. "Broad Trends Identified by the 2005 Patient Safety Survey" [accessed on May 21, 2007, July/August 2006]. Available at http://www.va.gov/ncps/TIPS/Docs/TIPS_JulAug06.pdf

Morey, J. C., R. Simon, G. D. Jay, R. L. Wears, M. Salisbury, K. A. Dukes, and S. D. Berns. 2002. "Error Reduction and Performance Improvement in the Emer-

gency Department through Formal Teamwork Training: Evaluation Results of the MedTeams Project." *Health Service Research* 37 (6): 1553–81.

Pronovost, P. J., T. Nolan, S. Zeger, M. Miller, and H. Rubin. 2004. "How Can Clinicians Measure Safety and Quality in Acute Care?" *Lancet* 363 (9414): 1061–7.

Pronovost, P. J., B. Weast, C. G. Holzmueller, B. J. Rosenstein, R. P. Kidwell, K. B. Haller, E. R. Feroli, J. B. Sexton, and H. R. Rubin. 2003. "Evaluation of the Culture of Safety: Survey of Clinicians and Managers in an Academic Medical Center." *Quality and Safety in Health Care* 12 (6): 405–10.

Rall, M., and P. Dieckmann. 2005. "Safety Culture and Crisis Resource Management in Airway Management: General Principles to Enhance Patient Safety in Critical Airway Situations." *Best Practise and Research Clinical Anaesthesiology* 19 (4): 539–57.

Reason, J. 2000. "Human Error: Models and Management." *Western Journal of Medicine* 172 (6): 393–6.

Roberts, K. H. 1990. "Some Characteristics of One Type of High Reliability Organization." *Organization Science* 1 (2): 160–77.

Roberts, K. H., and C. Tadmor. 2002. "Lessons Learned from Non-Medical Industries: The Tragedy of the USS Greeneville." *Quality and Safety in Health Care* 11 (4): 355–57.

Romano, P. S., J. J. Geppert, S. Davies, M. R. Miller, A. Elixhauser, and K. M. Mc-Donald. 2003. "A National Profile of Patient Safety in U.S. Hospitals." *Health Affairs (Millwood)* 22 (2): 154–66.

Rosen, A. K., D. M. Gaba, M. Meterko, P. Shokeen, S. J. Singer, S. Zhao, A. Labonte, and A. Falwell. In Press. "Recruitment of Hospitals for a Safety Climate Study: Facilitators and Barriers." *Joint Commission Journal on Quality and Patient Safety*.

Sexton, J., R. Helmreich, T. Neilands, K. Rowan, K. Vella, J. Boyden, P. Roberts, and E. Thomas. 2006a. "The Safety Attitudes Questionnaire: Psychometric Properties, Benchmarking Data, and Emerging Research." *BMC Health Services Research* 6: 44.

Sexton, J. B., M. A. Makary, A. R. Tersigni, D. Pryor, A. Hendrich, E. J. Thomas, C. G. Holzmueller, A. P. Knight, Y. Wu, and P. J. Pronovost. 2006b. "Teamwork in the Operating Room: Frontline Perspectives among Hospitals and Operating Room Personnel." *Anesthesiology* 105 (5): 877–84.

Sexton, J. B., E. J. Thomas, and R. L. Helmreich. 2000a. "Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys." *British Medical Journal* 320 (7237): 745–49.

———. 2000b. "Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys." *British Medical Journal* 320 (7237): 745–49.

Shekelle, P. G., R. L. Kravitz, J. Beart, M. Marger, M. Wang, and M. Lee. 2000. "Are Nonspecific Practice Guidelines Potentially Harmful?" *Health Services Research* 34 (7): 1429–48.

Shojania, K. G., K. M. McDonald, R. M. Wachter, and D. K. Owens. 2001. "Closing the Quality Gap: A Critical Analysis of Quality Improvement Strategies." *Medical Care* 29 (8): 709–26.

Shortell, S. M., J. E. Zimmerman, D. M. Rousseau, R. R. Gillies, D. P. Wagner, E. A. Draper, W. A. Knaus, and J. Duffy. 1994. "The Performance of Intensive Care Units: Does Good Management Make a Difference?" *Medical Care* 32 (5): 508–25.

Shortell, S. M., D. M. Rousseau, R. R. Gillies, K. J. Devers, and T. L. Simons. 1991. "Organizational Assessment in Intensive Care Units (ICUs): Construct Development, Reliability, and Validity of the ICU Nurse-Physician Questionnaire." *Medical Care* 29 (8): 709–26.

Singer, S. J. 2007. "Safety Climate in US Hospitals: Its Measurement, Variation, and Relationship to Organizational Safety Performance" Ph.D. Program in Health Policy. Cambridge, MA: Harvard University.

Singer, S. J., D. M. Gaba, J. J. Geppert, A. D. Sinaiko, S. K. Howard, and K. C. Park. 2003. "The Culture of Safety: Results of an Organization-Wide Survey in 15 California Hospitals." *Quality and Safety in Health Care* 12 (2): 112–8.

Singer, S. J., M. Meterko, L. Baker, D. Gaba, A. Falwell, and A. Rosen. 2007. "Workforce Perceptions of Hospital Safety Culture: Development and Validation of the Patient Safety Climate in Healthcare Organizations Survey." *Health Services Research* 42 (5): 1999–2021.

Sorra, J. N. V., T. Famolaro, and N. Dyer. 2007. *Hospital Survey on Patient Safety Culture: 2007 Comparative Database Report.* Rockville, MD: Agency for Healthcare Research and Quality.

Tucker, A., and A. Edmondson. 2003. "Why Hospitals Don't Learn from Failures: Organizational and Psychological Dynamics That Inhibit System Change." *California Management Review* 45: 55–72.

Ware, J. E., W. J. Harris, B. Gandek, B. W. Rogers, and P. R. Reese. 1997. *Multitrait/Multi-Method Analysis Program—Revised.* Boston: Health Assessment Lab.

Weick, K. E., and K. M. Sutcliffe. 2001. *Managing the Unexpected: Assuring High Performance in an Age of Complexity.* San Francisco: Jossey-Bass.

———. 2003. "Hospitals as Cultures of Entrapment: A Re-Analysis of The Bristol Royal Infirmary." *California Management Review* 45 (2): 73.

Weingart, S. N., A. N. Ship, and M. D. Aronson. 2000. "Confidential Clinician-Reported Surveillance of Adverse Events among Medical Inpatients." *Journal of General Internal Medicine* 15 (7): 470–7.

Welsh, C. H., R. Pedot, and R. J. Anderson. 1996. "Use of Morning Report to Enhance Adverse Event Detection." *Journal of General Internal Medicine* 11 (8): 454–60.

Zohar, D. 2002. "Modifying Supervisory Practices to Improve Subunit Safety: A Leadership-Based Intervention Model." *Journal of Applied Psychology* 87 (1): 156–63.

Supplementary Material

The following supplementary material for this article is available online:
Appendix SA1: HSR Author Matrix.

This material is available as part of the online article from http://www. blackwell-synergy.com/doi/abs/10.1111/j.1475-6773.2008.00839.x (this link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.