**Journal of Perinatology (2006) 26**, 463–470

www.nature.com/jp

ORIGINAL ARTICLE

# Variation in caregiver perceptions of teamwork climate in labor and delivery units

**JB Sexton[1,2], CG Holzmueller[1], PJ Pronovost[1,2,3], EJ Thomas[4], S McFerran[5], J Nunes[5], DA Thompson[1], AP Knight[6], DH Penning[1,7] and HE Fox[7]**

*[1]Department of Anesthesiology and Critical Care Medicine, The Johns Hopkins University, Baltimore, MD, USA; [2]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; [3]Department of Surgery, The Johns Hopkins University, Baltimore, MD, USA; [4]The Department of Medicine and Center of Excellence for Patient Safety at The University of Texas – Houston Medical School, Houston, TX, USA; [5]Northern California Regional Risk Management, Kaiser Permanente, Oakland, CA, USA; [6]Department of Management, The Wharton School, University of Pennsylvania, PA, USA and [7]Department of Obstetrics and Gynecology, The Johns Hopkins University, Baltimore, MD, USA*

**Objective:** To test the psychometric soundness of a teamwork climate survey in labor and delivery, examine differences in perceptions of teamwork, and provide benchmarking data.

**Design:** Cross-sectional survey of labor and delivery caregivers in 44 hospitals in diverse regions of the US, using the Safety Attitudes Questionnaire teamwork climate scale.

**Results:** The response rate was 72% (3382 of 4700). The teamwork climate scale had good internal reliability (overall $\alpha = 0.78$). Teamwork climate scale factor structure was confirmed using multilevel confirmatory factor analyses (CFI = 0.95, TLI = 0.92, RMSEA = 0.12, $SRMR_{within} = 0.04$, $SRMR_{between} = 0.09$). Aggregation of individual-level responses to the L&D unit-level was supported by ICC (1) = 0.06 ($P<0.001$), ICC (2) = 0.83 and mean $r_{wg(j)} = 0.83$. ANOVA demonstrated differences between caregivers $F$ (7, 3013) = 10.30, $P<0.001$ and labor and delivery units, $F$ (43, 1022) = 3.49, $P<0.001$. Convergent validity of the scale scores was measured by correlations with external teamwork-related items: collaborative decision making ($r = 0.780$, $P<0.001$), use of briefings ($r = 0.496$, $P<0.001$) and perceived adequacy of staffing levels ($r = 0.593$, $P<0.001$).

**Conclusion:** We demonstrate a psychometrically sound teamwork climate scale, correlate it to external teamwork-related items, and provide labor and delivery teamwork benchmarks. Further teamwork climate research should explore the links to clinical and operational outcomes. *Journal of Perinatology* (2006) **26,** 463–470. doi:10.1038/sj.jp.7211556; published online 15 June 2006

**Keywords:** teamwork; culture; climate collaboration; psychometrics; scale

Correspondence: Dr JB Sexton, Quality & Safety Research Group, 1909 Thames Street, 2nd floor, Baltimore, MD 21231, USA.
E-mail: jsexton2@jhmi.edu
Received 5 May 2006; accepted 22 May 2006; published online 15 June 2006

## Introduction

After the release of *To Err is Human*,[1] many healthcare institutions and organizations began the process of moving safety improvement efforts forward. One principle in this report was the 'promotion of effective team functioning.' The American College of Obstetricians and Gynecologists (ACOG) responded with a set of safety-related objectives for clinical providers to follow in daily practice.[2] One objective stressed commitment to a patient safety culture through the daily practice of teamwork, communication, collaboration and strong leadership for providers.

Effective teamwork is critical in high-risk settings where individuals interact with other persons to perform their job. In aviation, plane crashes resulting from flight crew discord prompted development of crew resource management (CRM) training to address team climate and improve performance.[3,4] In healthcare, researchers have identified and are investigating group hierarchy, stressful work environments, poor communication and varying perceptions of what comprises a team as some barriers to effective teamwork.[5,6]

One outcome of poor team climate is medical error. The Joint Commission on Accreditation of Healthcare Organization's (JCAHO) sentinel event investigation in labor and delivery (L&D) found poor communication as a root cause in over 80% of perinatal deaths and injuries.[7] Additionally, the leading root cause of perinatal deaths and injuries tracked by JCAHO was communication breakdowns, which was cited in over 80% of events.[8] In another study, poor teamwork was attributed to 40% of maternal deaths and 45% of near miss morbidities.[9] Team performance is important in L&D because a normal situation can transition to an emergency rather quickly. A rescue team must assemble quickly, communicate clearly and collaborate effectively to avoid needless morbidity or mortality.[10,11]

Organizational culture also plays a major role in guiding individual behaviors and ultimately team performance.[12–14] Recently, JCAHO announced that hospitals will be required to measure culture annually, possibly in 2007 (www.jcaho.org). Aviation has measured culture for decades, initially with the Cockpit Management Attitudes Questionnaire (CMAQ)[15] and its successor the Flight Management Attitudes Questionnaire (FMAQ),[16] both of which are reliable, sensitive to change[17] and the attitudes elicited shown to predict performance.[18,19] There is emerging evidence that measurements of attitudes in healthcare settings are also sensitive to interventions.[20–22] (Martin Makary, MD, preliminary data, August, 2005) Recent evidence from the operating room demonstrated that caregiver assessments of teamwork climate can be reliably captured through a psychometrically sound survey instrument.[23]

The primary aims of this study were to test the reliability of a teamwork climate scale in L&D and provide teamwork climate benchmarking data for L&D units and caregiver types. Secondary objectives were to examine differences in perceptions of teamwork by provider type and L&D unit.

## Methods

*Climate vs culture: a clarification of terms*

Organizational researchers view the distinction between climate and culture as the difference between taking a snapshot during one brief time period (measuring climate), versus measuring the underlying determinants of climate (culture).[24] Whereby climate can be captured by survey instruments, culture is better understood through ethnographic and anthropologic studies of the artifacts,[25,26] values,[27] and assumptions[26,28] that make an organization distinct.[29] Metaphorically, culture is a complex grid of interconnected highway systems, while climate is the traffic that maneuvers the streets. Climate is easier to measure and influence than the deeper culture. Technically, the distinction between climate and culture comes down to this: if you measure it with surveys and intend to publish it, you are measuring *climate*. That said, administrators, frontline caregivers, researchers and even funding agencies have taken to calling this area of inquiry in healthcare 'safety culture research.' Hence, we will use safety culture to refer to the larger endeavor and save *climate* for discussing the results from the survey.

*Design and study population*

The Safety Attitudes Questionnaire (SAQ) is a psychometrically sound inventory of frontline caregiver assessments regarding the work environment and the context in which they deliver care.[30,31] The SAQ was refined from the Intensive Care Unit Management Attitudes Questionnaire (ICUMAQ),[32] which was adapted from aviation.[15,16] After reviewing the literature and conducting roundtable discussions with L&D caregivers, we found the content appropriate for L&D and did not identify any new items.

The SAQ was administered between October 2002 and October 2004 to sixteen L&D caregiver types in 44 hospitals in the northeast, mid-atlantic and west coast regions of the US. Hospitals varied in size and teaching status. To maximize comparability, we restricted analyses to L&D caregiver types most common to all hospitals; including obstetricians, pediatricians, anesthesiologists, certified registered nurse anesthetists (CRNA), registered nurses (RN), licensed vocational nurses (LVN)/OB technicians, nurse managers/charge nurses and combined perinatologists and neonatologists. Surveys that were blank or had invariant responses (e.g., all responses were 'agree strongly') were excluded from this analysis, as they did not provide any diagnostic information. Surveys were administered during pre-existing departmental and staff meetings with a pencil and sealable return envelope. Individuals not captured in pre-existing meetings, were hand delivered a survey, pencil and return envelope. A local physician champion at each site assisted with hand delivery of surveys to physicians absent during these meetings. We have found this administration technique to generate relatively high response rates.[22]

*Measurements*

The SAQ measures six domains; including, teamwork climate, safety climate, job satisfaction, perceptions of management, stress recognition and working conditions. Here, we report results from the teamwork climate domain. Teamwork climate assesses how healthcare providers from the same work unit perceive the quality of collaboration between personnel in that unit. Six of the 30 SAQ scale items define the teamwork climate domain, with the response scale ranging from 1 (disagree strongly) to 5 (agree strongly).

*Statistical analysis*

To verify the single factor nature of the teamwork climate scale, we performed a multilevel confirmatory factor analysis (MCFA) to account for the nesting of individual caregivers within L&D units. The MCFA corrects the between group covariance matrix so that an unbiased between group factor structure is obtained.[33,34] Using MPLUS code specified by Dyer *et al*.,[34] we performed MCFAs on the 6-item teamwork climate scale.

A basic criterion required to adequately assess culture or climate constructs is that individual perceptions show high agreement within units (e.g., L&D units) and high variance between units.[35] Teamwork climate is conceptualized at the L&D unit level of analysis, so we calculated intraclass correlation coefficients (ICCs) and the $r_{wg(j)}$ statistic to justify aggregation of caregivers within their L&D units. It is important to establish whether L&D units differ in their teamwork climate scores. To justify aggregation, it is necessary to first establish that caregivers within an L&D unit have similar perceptions of the units teamwork climate (i.e., that there is little variance within units in perceptions of teamwork climate). The $r_{wg(j)}$ statistic is a measure of consensus, typically ranging

from 0 to 1 that provides a useful index of within-group agreement.[36] A second prerequisite to aggregation is establishing that L&D units differ in their teamwork climate scores (i.e., that there is variance across units in perceptions of teamwork climate). The ICC (1) statistic is a measure of between group variability and the ICC (2) statistic is a measure of the reliability of the group means.[37] To calculate ICC (1) and ICC (2), a one-way analysis of variance is conducted on the individual level responses, with L&D unit as the independent variable.

We then used reliability analyses to evaluate the 6-item teamwork climate scale. Internal reliability was assessed using Cronbach's $\alpha$.[38] To test for differences in perceptions of teamwork, we focused on teamwork climate by caregiver and by hospital. To improve interpretability and representativeness, we analyzed groups of caregivers with over 100 respondents, and excluded smaller caregiver groups. In the case of perinatologists ($n = 65$) and neonatologists ($n = 82$), we collapsed the two distinct specialties into one caregiver category after MANOVA revealed no significant differences in responses to any of the teamwork items $F$ (6, 136) = 0.636, $P < 0.701$. Using MANOVA, we tested for differences between provider types and differences between hospitals with respect to each teamwork climate item. We then used ANOVA to test the same groups for differences on the teamwork climate scale score. Teamwork climate scale scores were computed by taking the mean of the six items (one item was reverse scored due to the negative wording). In addition to the means used in MANOVA, ANOVA and internal scale reliabilities, we also report the percent agreement (agree slightly plus agree strongly) for items and scale scores of each position and hospital. We call this 'percentage agree' or 'percentage reporting good teamwork.' In exploratory analyses to put L&D teamwork climate into context at the unit level, we correlated mean L&D unit teamwork climate scores with mean unit item scores from teamwork related SAQ items using exploratory two-tailed Pearson correlations. All statistical analyses were performed using SPSS version 13.0 (Chicago, IL, USA), SAS version 9.1 and MPLUS version 2.01.

## Results

### *Respondent demographics*

Of 3395 returned surveys, 13 were excluded because of invariant data (e.g. all strong agree responses) or no responses. There were 3382 usable L&D respondents from the 44 hospitals studied. Overall response rate was 72% (3382 out of 4700), with a range across hospitals of 42 to 100%. Registered nurses accounted for 56% ($n = 1877$) of respondents, obstetricians 15% ($n = 494$), LVN/OB technicians 7% ($n = 227$), anesthesiologists 6% ($n = 213$) and there was 4% each for perinatologists and neonatologists ($n = 147$), pediatricians ($n = 144$), CRNAs ($n = 144$) and nurse managers/charge nurses ($n = 136$).

Table 1 lists response rates and respondent demographics. Mean age by caregiver was 44, with a consistent range from 42 to 46. Obstetricians had the most experience (mean 16 years ±9.7) and LVN/OB technicians had the least (mean 9 years ±8.6), with the latter having the lowest job turnover (9 years ±8 working at current hospital). Physicians were predominantly male with obstetricians having the highest percentage of females (50%) and nurse positions most often female, nurse managers 99% and RNs 97%.

### *Teamwork climate scale psychometrics*

Here, we describe how caregiver responses to six teamwork climate items were used to generate a representative and reliable teamwork climate score. First, and consistent with prior development work, the 6-item teamwork scale exhibited acceptable internal consistency reliability (overall $\alpha = 0.78$; obstetrician $\alpha = 0.78$; perinatologists and neonatologists $\alpha = 0.79$; pediatrician $\alpha = 0.82$; anesthesiologist $\alpha = 0.82$; CRNA $\alpha = 0.89$; registered nurse $\alpha = 0.75$; LVN/OB technician $\alpha = 0.70$ and nurse manager/charge nurse $\alpha = 0.81$). Second, a multilevel confirmatory factor analysis was conducted to validate the single, multilevel factor structure of the teamwork climate construct at the L&D unit level. A single, latent multilevel structure fit the data well (CFI = 0.95, TLI = 0.92, RMSEA = 0.12, $SRMR_{within} = 0.04$,

**Table 1** Respondent demographics

| *Position* | *Response rate (returned/admin)* | *Age (mean years ± s.d.)* | *% Female (number)* | *Years experience in position (mean ± s.d.)* | *Years working at current hospital (mean ± s.d.)* |
|---|---|---|---|---|---|
| Obstetrician | 67% (494/739) | 45 (9.91) | 50 | 16 (9.67) | 11 (9.36) |
| Peri & neonatologist | 77% (147/192) | 45 (9.31) | 44 | 14 (9.26) | 11 (10.33) |
| Anesthesiologist | 54% (213/401) | 44 (7.83) | 26 | 13 (7.35) | 8 (6.84) |
| Pediatrician | 54% (144/266) | 43 (9.71) | 47 | 12 (8.10) | 9 (7.11) |
| Registered nurse | 77% (1877/2442) | 42 (10.71) | 97 | 12 (9.61) | 9 (8.46) |
| Nurse manager/charge nurse | 79% (136/172) | 46 (7.56) | 99 | 18 (8.11) | 13 (8.05) |
| CRNA | 69% (144/208) | 46 (8.98) | 63 | 14 (9.75) | 9 (7.65) |
| LVN/OB technicians | 81% (227/280) | 42 (10.95) | 91 | 9 (8.61) | 9 (8.07) |

$SRMR_{between} = 0.09$) and all items had standardized loadings greater than .65. Overall, these results provide support for the reliability and validity of the 6-item scale as an assessment of teamwork climate at the hospital L&D unit level.

To further examine the legitimacy of aggregating individual-level responses to the L&D unit level (i.e., L&D caregiver perceptions of teamwork within a hospital cluster together consistently, indicating specificity to L&D in that particular hospital), we examined a number of statistical indices. First, the variance in perceptions of teamwork climate accounted for by group membership was in the range of acceptable, $ICC(1) = 0.06$, $P<0.001$. Second, the reliability of the group mean for the construct was also within the range of acceptable values $ICC(2) = 0.83$. Third, overall within-unit agreement was greater than the standard 0.70 threshold, as evidenced by an average $r_{wg(j)}$ across units of 0.83. In this sample, $r_{wg(j)}$ values ranged from 0.53 to 0.91, with a s.d. of 0.07. Together, these indices provide good justification for the legitimacy of L&D teamwork climate as a collective-level construct.

*Variation in teamwork climate by caregiver and hospital*

MANOVA of the six items yielded two significant omnibus $F$ results. An omnibus F for L&D caregivers of $F$ (42, 4614) = 2.16, $P<0.001$, indicating that L&D caregivers perceive specific teamwork issues differently as a function of their role. The teamwork climate items demonstrated differences, particularly in heeding nurse input, physician–nurse collaboration, conflict resolution and ease in asking questions. An omnibus $F$ for hospitals of $F$ (138, 5740) = 1.48, $P<0.001$ indicates that respondents perceive teamwork issues differently as a function of the hospital in which they work.

After testing for differences on the teamwork climate *items*, we tested for differences on the *scale scores*. ANOVA demonstrated significant differences in teamwork climate scale scores between caregivers $F$ (7, 3013) = 10.30, $P<0.001$ and between L&D units, $F$ (43, 1022) = 3.49, $P<0.001$. Figure 1 shows the percent agreement (agree slightly and agree strongly) by L&D caregiver and by hospital.

Table 2 displays each teamwork climate item and descriptive statistics for benchmarking teamwork climate data for L&D caregivers. Overall, physicians and nurse managers perceived teamwork issues more positively than nurses and LVN/OB Techs. Eighty-four percent of obstetricians felt nurse input was well received, compared to 50% of CRNAs and 61% of RNs. Seventy-percent of perinatologists and neonatologists disagreed with the statement 'in this clinical area it is difficult to speak up if I perceive a problem with patient care,' whereas only 51% of RN and LVN/OB technicians each disagreed. Perinatologists and neonatologists consistently perceived appropriate resolution of disagreements (74% agreed) and felt physicians and nurses work as a well coordinated team (84% agreed), whereas CRNAs scored lowest for each item, at 44 and 50%, respectively.

*Exploring the content and convergent validity of the teamwork climate scale*

In exploratory analyses we tested the content and convergent validity of the teamwork climate scale relative to 10 teamwork-related items (Table 3). Individual caregiver responses were aggregated from the entire unit to compare unit level teamwork climate scale scores to unit level item results. L&D unit teamwork climate was positively correlated with collaborative decision making ($r = 0.780$, $P<0.001$), briefing personnel before a procedure ($r = 0.496$, $P<0.001$) and communication of issues at shift change ($r = 0.496$, $P<0.001$). With regard to perceptions of how busy the unit is, L&D teamwork climate was not significantly related to perceptions of workload on the unit ($r = -0.259$, $P = 0.090$), but was significantly related to the perceived adequacy of staffing levels ($r = 0.593$, $P<0.001$). In addition, familiarity with colleagues (e.g. knowing their names) was positively associated with teamwork climate ($r = 0.473$, $P = 0.019$).

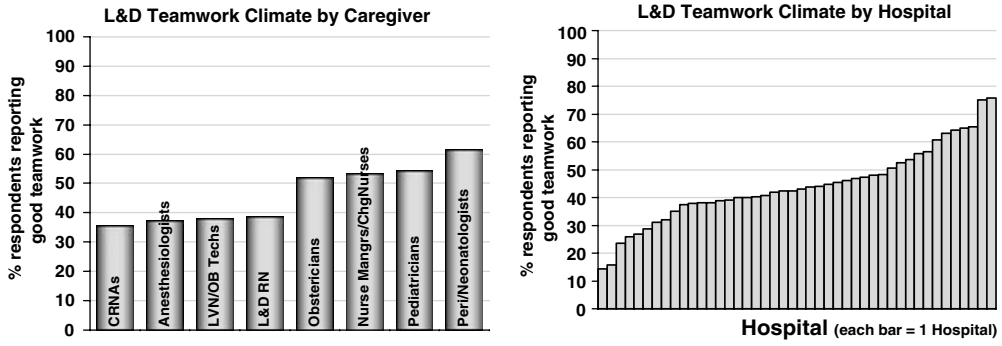


**Figure 1** shows the percent agreement (agree slightly and agree strongly) by OR caregiver and by hospital (each hospital is a distinct L&D unit).

**Table 2** Teamwork climate score by caregiver type

| *Teamwork climate scale Item* | *Overall mean(s.d.)*[a] *% agree (min–max)* | *OB mean(s.d.)*[a] *% agree (min–max)* | *Peri and Neonat mean(s.d.)*[a] *% agree (min–max)* | *Ped mean(s.d.)*[a] *% agree (min–max)* | *Anesth mean(s.d.)*[a] *% agree (min–max)* | *CRNAs mean(s.d.)*[a] *% agree (min–max)* | *R.N. mean(s.d.)*[a] *% agree (min–max)* | *LVN/OB Tech mean(s.d.)*[a] *% agree (min–max)* | *Nurse Mgr/chrge mean(s.d.)*[a] *% agree (min–max)* |
|---|---|---|---|---|---|---|---|---|---|
| I have the support I need from other personnel to care for our patients | **3.91(0.968)** *74.7* (45.3–100.0) | **3.98(0.927)** *77.8* (50.0–100.0) | **4.02(0.903)** *81.6* (60.0–100.0) | **3.93(0.901)** *75.4* (50.0–100.0) | **3.67(0.985)** *67.5* (20.0–100.0) | **3.64(1.180)** *64.6* (40.0–100.0) | **3.92(0.964)** *75.2* (36.5–100.0) | **3.83(0.988)** *67.5* (20.0–100.0) | **4.15(0.893)** *83.8* (62.5–100.00) |
| It is easy for personnel in this clinical area to ask questions when there is something that they do not understand | **4.10(0.921)** *80.1* (66.0–100.0) | **4.07(0.841)** *79.1* (50.0–100.0) | **4.27(0.761)** *87.8* (71.4–100.0) | **4.03(0.773)** *77.5* (50.0–100.0) | **3.72(1.069)** *67.5* (12.5–100.0) | **3.66(1.178)** *62.5* (33.3–92.3) | **4.15(0.914)** *82.4* (65.1–100.0) | **4.09(0.880)** *78.0* (41.7–100.0) | **4.38(0.789)** *89.0* (62.5–100.0) |
| Nurse input is well received in this clinical area | **3.76(1.074)** *66.9* (28.6–87.9) | **4.17(0.863)** *84.5* (60.0–100.0) | **4.32(0.802)** *83.7* (57.1–100.0) | **4.22(0.829)** *83.7* (63.6–100.0) | **3.76(1.002)** *67.6* (0–100.0) | **3.44(1.170)** *50.0* (14.3–100.0) | **3.57(1.117)** *60.8* (31.3–92.3) | **3.78(0.942)** *62.6* (20.0–84.6) | **4.05(0.984)** *75.0* (55.6–100.0) |
| In this clinical area, it is difficult to speak up if I perceive a problem with patient care (**reverse scored**)[b] | **3.49(1.202)** *55.4* (27.7–81.8) | **3.74(1.121)** *62.4* (20.100.0) | **3.97(1.044)** *69.7* (40.0–100.0) | **3.61(1.040)** *61.0* (42.9–90.0) | **3.62(1.167)** *60.2* (25.0–100.0) | **3.57(1.231)** *60.1* (12.5–100.0) | **3.36(1.213)** *51.2* (11.8–77.8) | **3.44(1.246)** *51.3* (16.7–85.7) | **3.68(1.286)** *61.8* (44.4–100.0) |
| Disagreements in this clinical area are appropriately resolved (i.e., not who is right, but what is best for the patient) | **3.36(1.114)** *49.5* (17.6–70.4) | **3.54(1.067)** *55.4* (25.0–81.8) | **3.84(0.944)** *74.3* (40.0–87.0) | **3.71(0.906)** *61.6* (33.3–77.8) | **3.30(1.020)** *48.4* (20.0–85.7) | **3.21(1.251)** *43.8* (16.0–88.9) | **3.24(1.122)** *44.8* (9.8–72.7) | **3.43(1.132)** *49.8* (20.0–100.0) | **3.64(1.140)** *61.0* (0–100.0) |
| The doctors and nurses here work together as a well coordinated team | **3.74(1.052)** *68.0* (21.4–97.0) | **3.92(0.943)** *76.5* (50.0–100.0) | **4.09(0.891)** *83.7* (54.5–100.0) | **4.00(0.955)** *79.6* (63.6–100.0) | **3.33(1.159)** *55.2* (28.6–100.0) | **3.42(1.132)** *50.0* (11.1–100.0) | **3.71(1.051)** *66.6* (15.4–91.7) | **3.74(1.021)** *64.8* (20.0–100.0) | **3.79(1.176)** *71.3* (37.5–100.0) |

[a]Mean score on a 0–5 scale (5 = agree strongly).
[b]Item is reverse scored, respondent answer in the positive direction was to 'disagree.' Percent disagreement was reported here for this item.

**Table 3** Nonscaled items used in unit level correlations to measure convergent validity of the teamwork climate scale

| *Nonscaled teamwork-related item* | *2-Tailed pearson/P-value* |
|---|---|
| Briefings of personnel before a procedure (e.g., intubation, central venous line) are common in this clinical area. | $r = 0.496$ $P < 0.001$ |
| High levels of workload are common in this clinical area. | $r = -0.259$ $P = 0.090$ |
| The levels of staffing in this clinical area are sufficient to handle the number of patients. | $r = 0.593$ $P < 0.001$ |
| Decision making in this clinical area utilizes input from relevant personnel. | $r = 0.780$ $P < 0.001$ |
| During emergencies, I can predict what other personnel are going to do next. | $r = 0.341$ $P = 0.023$ |
| I am frequently unable to express disagreement with staff physicians. | $r = -0.363$ $P = 0.015$ |
| I know the first and last names of all the personnel I worked with during my last shift. | $r = 0.473$ $P = 0.019$ |
| I feel burned out from my work. | $r = -0.407$ $P = 0.019$ |
| Important issues are well communicated at shift changes. | $r = 0.496$ $P < 0.001$ |
| Communication breakdowns which lead to delays in starting surgical procedures are common | $r = -0.381$ $P = 0.011$ |

## Discussion

In labor and delivery, perceptions of teamwork climate are affected by the environment and role within the team. Teamwork climate is perceived differently as a function of the L&D unit in which a caregiver works, and L&D caregivers have discrepant attitudes about teamwork as a function of their role. These results are not unlike previous findings in the operating room, where the teamwork climate scale was psychometrically sound and reliable across caregiver types.[23] The teamwork climate scale is a composite measure of the extent to which caregivers report that they feel supported, can speak up comfortably, can ask questions, feel nurse input is heeded, that conflicts are resolved, and that physicians and nurses collaborate. Here, we reported how the scale detects differences in perceptions of teamwork as a function of caregiver type and as a function of the L&D unit in which the respondent delivers care.

Correlations between the teamwork climate scale and the 10 teamwork-related items further illustrate and define issues related to teamwork climate. In addition, the content and interpretable relationship of these 10 items to the teamwork climate scale provides convergent validity, demonstrating that the teamwork climate scale is indeed eliciting perceptions of collaboration in labor and delivery.

The teamwork climate scale scores at the L&D unit level were associated with better information management at point of care transitions like communication during shift changes and briefings before a procedure. Good teamwork climate, as measured by the scale, also appears to be related to familiarity with other caregivers, such that knowing the names of ones' colleagues and being able to predict their actions during emergencies was associated with unit-level teamwork climate. In L&D units where caregivers report good teamwork climate scale scores, they also report that decisions are made collaboratively and, when necessary, it is possible to disagree with staff physicians. Good teamwork climate scale scores were associated with lower levels of caregiver burnout from their work – which has tremendous implications for using teamwork training to combat the epidemic of nurse retention. In another operationally relevant way, poor teamwork climate was associated with communication breakdowns that led to delays, which are very costly and unproductive periods for hospitals. Finally, the relationship between the teamwork climate scale score and perceptions of being busy suggested that teamwork climate is related more to perceptions of *adequate staffing levels* and less to perceptions of *workload*. It may be that good teamwork climate creates a synergy that offsets perceptions of inadequate staffing levels independent of perceptions of workload.

Overall, physicians and nurse managers were much more satisfied than nurses, with respect to the collaboration they experienced. The global difference between nurses and physicians has been documented in other clinical areas[32,39] and may be due primarily to personal characteristics of caregivers. A suggestion repeated in the comment section of the SAQ was to 'improve collaboration between physicians and nurses.' Thomas and co-workers found that provider characteristics (personal attributes, reputation and expertise/seniority) influenced the ability of neonatal intensive care unit caregivers to work together.[6]

In comparing CRNAs to anesthesiologists, we found comparable scores in five of the six teamwork climate items, whereby anesthesiologists were $\leq 5$ percentage points higher than CRNAs. This comparability may be due to similar clinical roles in labor and delivery. Interestingly, there was a significant response difference for the item 'nurse input is well received in this clinical area,' with half of CRNAs agreeing relative to two thirds of the anesthesiologists.

In addition, perinatologists and neonatologists reported more agreement, 70 to 88% with all six teamwork climate items while CRNAs and RNs demonstrated generally lower levels of agreement. This difference may be a function of job roles in the labor and delivery process. CRNAs interface with the team to provide pain relief or anesthetic for the patient, often in crisis mode. Perinatologists and neonatologists are maternal-fetal medicine specialists, and may be a step removed from the routine L&D team

process since they care for special needs babies and may experience better L&D teamwork due to the specialized nature of their interactions with other caregivers. Labor and delivery teams are extremely complex with as many as 16 caregiver types collaboratively and independently providing care. It will be important to look at group dynamics and sub-group dynamics among L&D caregivers in the future.

While the L&D respondents overall scored high for the item 'it is easy for personnel in this clinical area to ask questions…' (80% agreed), only 55% found it easy to speak up if they perceived a problem with patient care, and only half felt that conflicts were appropriately resolved. This disconnect may be a by-product of the autonomy hierarchy seen in many industries, such as medicine and aviation.[40] In medicine, questions seeking advice or knowledge are welcome, whereas questioning someone's performance or disagreeing with their actions is taboo. Indeed, suggestions for improving patient safety in their L&D unit from the open-ended comments section of the survey frequently noted better communication among caregivers, between units and from management.

We recognize several limitations to our study. First, we included data from only 44 hospitals. In addition, our sample includes academic, faith based, and community hospitals from the northeast, mid-atlantic and west coast regions of the US. As such, we do not yet know whether these results are generalizable. Nevertheless, the 15 most recent units that we added to the sample fell within the existing range of teamwork climate scores, indicating stability in the distribution of scores. Also, all 44 hospitals administered the SAQ prior to teamwork interventions, giving us a baseline distribution of scores. Second, the criterion validity of teamwork climate in L&D units has yet to be established. As such, we cannot say if L&D teamwork climate is related to, for example, actual delays, annual nurse turnover, or error rates. Establishing the criterion validity of L&D teamwork climate is a critical next step in this research. Third, the data reported from these 44 units do not show whether teamwork climate is sensitive to change after exposure to training or an intervention. There is emerging evidence from operating room studies demonstrating that teamwork climate can be improved, and these improvements are associated with impressive clinical and operational improvements.[20] (Martin Makary, MD unpublished data, August 2005). We look forward to future research to further evaluate the psychometric soundness of teamwork climate and to investigate the clinical and operational variables associated with teamwork climate.

JCAHO's recent recognition of culture as an important component in hospital safety has prompted many to seek scientific methods to measure culture. The SAQ elicits caregiver attitudes that can assess culture in L&D units. Specifically, we demonstrated how the interpersonal component of the L&D work environment can be measured using the teamwork climate scale of the SAQ. Furthermore, we demonstrated that the teamwork climate measure of the SAQ assesses a unit-level construct, such that a pattern of consensus is seen in caregiver responses regarding teamwork climate within their particular unit, and caregiver assessments vary across units. This psychometrically sound assessment provides benchmarks for obstetric and gynecologic departments and hospitals seeking to compare their teamwork climate to national means, and can serve as a baseline measure for evaluating interventions. The significant variation in L&D teamwork scores by hospital suggests that strategies to improve teamwork attitudes at some centers may be effective. Identification and dissemination of these best practices may also benefit the larger obstetric and gynecologic community.

Findings from this research add detail to teamwork research in healthcare[32,39] by providing a reliable teamwork climate scale, correlating it to external teamwork-related items, and providing an initial set of L&D teamwork benchmarks.

## Acknowledgments

This work was funded in part by Agency for Healthcare Research and Quality Grant # 1PO1HS1154401

## References

1 Kohn L, Corrigan J, Donaldson M. To err is human: building a safer health system. editors. Institute of Medicine Report . National Academy Press: Washington, DC 1999.
2 ACOG CoQIaPS. ACOG committee opinion: Patient safety in obstetrics and gynecology. *Int J Gynecol Obstet* 2004; **86**: 121–123.
3 Helmreich R. Managing human error in aviation. *Sci Am* 1997; **276**(5): 62–67.
4 Helmreich R, Wilhelm J. Outcomes of crew resource management training. *Int J Aviat Psychol* 1991; **1**: 287–300.
5 Thomas E, Sexton J, Helmreich R. Translating teamwork behaviors from aviation to healthcare: development of behavioral markers for neonatal resuscitation. *Qual Saf Health Care* 2004; **13**(1): i57–i64.
6 Thomas E, Sherwood G, Mulhollem J, Sexton BJ, Helmreich RL. Working together in the Neonatal Intensive Care Unit: Provider Perspectives. *J Perinatol* 2004; **24**: 552–559.
7 JCAHO. Preventing infant death and injury during delivery. *Sentinel Event Alert* 30. 7-21-2004. JCAHO.
8 JCAHO. Sentinel Event Alert #30: Preventing infant death and injury during deliver. JCAHO. 7-21-2004. http://www.jcaho.org/about+us/news+letters/sentinel+accessed 4-10-2005.
9 Geller S, Rosenberg B, Cox S. The continuum of maternal morbidity and mortality: factors associated with severity. *Am J Obstet Gynecol* 2004; **191**: 939–944.
10 Simpson K. Failure to rescue. Implications for evaluating quality of care during labor and birth. *J Perinat Neonat Nurs* 2005; **19**(1): 24–34.
11 Clarke S, Aiken L. Failure to rescue. *Am J Nurs* 2003; **103**(1): 42–47.
12 Schein E. Organizational culture. *Am Psychologist* 1990; **45**: 109–119.

13 Schneider B, Goldstein H, Smith D. The ASA framework: an update. *Personnel Psychol* 1995; **40**: 747–773.
14 Helmreich R, Merritt A. *Culture at Work in Aviation and Medicine: National, Organizational, and Professional Influences*. Ashgate Publishing Limited: Aldershot, UK, 1998 p 176.
15 Helmreich R. Cockpit management attitudes. *Human Factors* 1984; **26**: 583–589.
16 Helmreich R, Merrrit A, Sherman P, Gregorich S, Wiener E. The flight management attitudes questionnaire (FMAQ). *NASA/UT/FAA Technical Report 93–4*. Austin, TX: The University of Texas.
17 Gregorich S, Helmreich R, Wilhelm J. The structure of cockpit management attitudes. *J Appl Psychol* 1990; **75**: 682–690.
18 Helmreich R, Foushee H, Benson R, Russini W. Cockpit resource management: exploring the attitude-performance linkage. *Aviat Space Environ Med* 1986; **57**(12 Part 1): 1198–1200.
19 Foushee H. Dyads and triads at 25 000 feet: Factors affecting group process and aircrew performance. *Am Psychologist* 1984; **39**: 885–993.
20 DeFontes J, Surbida S. Preoperative safety briefing project. *Permanente J* 2004; **8**(2): 21–27.
21 Pronovost P, Weast B, Rosenstein B, Sexton JB, Holzmueller C, Paine L *et al.* Implementing and validating a comprehensive unit-based safety program. *J Patient Safety* 2005; **1**(1): 33–40.
22 Pronovost PJ, Sexton BJ. Assessing safety culture: guidelines and recommendations. *Qual Saf Health Care* 2005; **14**: 231–233.
23 Sexton J, Makary M, Tersigni A, Pryor D, Helmreich A, Thomas E *et al.* Teamwork in the operating room: Frontline perspectives among hospitals and operating room personnel. *Anesthesiology* 2006 in press.
24 Mearns K, Flin R, Fleming M, Gordon R. *Organization and Human Factors in Offshore Safety*. HSE: London, 1997.
25 Trice HM, Beyer JM. *The Cultures of Work Organizations*. Prentice Hall: Englewood Cliffs, NJ, 1993.
26 Schein EH. *Organizational Culture and Leadership: A Dynamic View*. Jossey-Bass: San Francisco, 1992.
27 Schwartz SH. Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. In: Zanna MP (ed). *Advances in Experimental Social Psychology*. Academic Press: San Diego, CA, 1992, pp 1–65.
28 Schein EH. Organizational culture. *Am Psychologist* 1990; **45**: 109–119.
29 Ostroff C, Kinicki AJ, Tamkins MM. Organizational culture and climate. In: Borman WC, Ilgen DR, Klimoski RJ (eds). *Handbook of Psychology*. Wiley & Sons: Hoboken, NJ, 2003, pp 565–594.
30 Thomas E, Sexton J, Neilands T, Frankel A, Helmreich R. The effect of executive walk rounds on nurse safety climate attitudes. A randomized trial of clinical units. *BMC Health Serv Res* 2005; **5**: 28 doi: 10.1186/1473-6963-5-28 [published online first: 11 April 2005].
31 Sexton J, Helmreich R, Neilands T, Rowan K, Vella K, Boyden J *et al.* The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res* 2006; **6**(44) doi:10.1186/1472-6963-6-44.
32 Thomas E, Sexton J, Helmreich R. Discrepant attitudes about teamwork among critical care nurses and physicians. *Crit Care Med* 2003; **31**(3): 956–959.
33 Muthen B. Multilevel covariance structure analysis. *Sociol Methods Res* 1994; **22**(Multilevel Modeling): 376–398.
34 Dyer N, Hanges P, Hall R. Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadership Quart* 2005; **16**(1): 149–167.
35 Klein KJ, Kozlowski SWJ. From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organ Res Meth* 2000; **3**: 211–236.
36 James L, Demaree R, Wolf G. Rwg: An assessment of within-group interrater agreement. *J Appl Psychol* 1984; **78**: 306–309.
37 Bliese P. Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In: Klein K, Kozlowski S (eds). *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*. Jossey-Bass: San Francisco, 2000, pp 349–381.
38 Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–333.
39 Wiles R. Teamwork in primary care: the views and experiences of nurses, midwives and health visitors. *J Advanced Nurs* 1994; **20**: 324–330.
40 Sexton J, Helmreich R, Thomas E. Error, stress and teamwork in medicine and aviation: Cross sectional surveys. *Br Med J* 2000; **320**: 745–749.