

Teamwork and Communication

Using the Communication and Teamwork Skills (CATS) Assessment to Measure Health Care Team Performance

Allan Frankel, M.D.
Roxane Gardner, M.D., M.P.H.
Laura Maynard, M.Div.
Andrea Kelly, R.N., J.D.

Effective teamwork and communication skills are cornerstones of safe, reliable, and high-quality health care. When embedded in the fabric of the daily routine, these skills heighten staff and patient satisfaction and facilitate optimal clinical outcomes.¹⁻³ When these behaviors are not robust, failures in teamwork and communication contribute to adverse medical events. Although nonmedical industries employ observations or audits to monitor and improve team processes and communication, there is no “gold standard” method for doing so in health care.^{4,5} Patient safety research is now targeting ways to observe and measure the teamwork skills of health care providers in a variety of high-intensity medical environments.⁶⁻⁹ Patient safety administrators, educators, and researchers are striving to understand how best to monitor and improve team skills and determine what approaches to monitoring best suit their organizations.

Review of closed malpractice claims and Joint Commission sentinel events consistently illustrates the important role of communication and teamwork in reducing and preventing harm to patients.¹⁰⁻¹³ Despite the valuable role these skills play in clinical care, team skills acquisition and proficiency has long been taken for granted. Medical and nursing educational curricula have generally not included formal training in teamwork. Consequently, most practicing health care professionals have little expertise with these skills.

True expertise is a function of knowledge, skills, and abilities not only in clinical matters but also in social mat-

Article-at-a-Glance

Background: Patient safety administrators, educators, and researchers are striving to understand how best to monitor and improve team skills and determine what approaches to monitoring best suit their organizations. A behavior-based tool, based on principles of crisis resource management (CRM) in nonmedical industries, was developed to quantitatively assess communication and team skills of health care providers in a variety of real and simulated clinical settings.

The CATS Assessment: The Communication and Teamwork Skills (CATS) Assessment has been developed through rapid-cycle improvement and piloted through observation of videotaped simulated clinical scenarios, real-time surgical procedures, and multidisciplinary rounds. Specific behavior markers are clustered into four categories—coordination, cooperation, situational awareness, and communication. Teams are scored in terms of the occurrence and quality of the behaviors. The CATS Assessment results enable clinicians to view a spectrum of scores—from the overall score for the categories to specific behaviors.

Conclusion: The CATS Assessment tool requires statistical validation and further study to determine if it reliably quantifies health care team performance. The patient safety community is invited to use and improve behavior-based observation measures to better evaluate their training programs, continue to research and improve observation methodology, and provide quantifiable, objective feedback to their clinicians and organizations.

ters, especially with respect to relationships with coworkers and our patients. We have produced clinicians with great expertise, but expert teamwork is lacking, and harm may come to patients because health care teams are not effectively managing clinical events. For example, teammates may lack a shared mental model of what needs to be done or how best to approach the task at hand, situation awareness may dim, resources may not arrive in a timely fashion, information may not be shared by all, briefings or handoffs may be incomplete or performed poorly, and teammates may be reluctant to speak up because of intimidation or fear.

Efforts are underway to develop and implement team-training programs for health care professionals and to create an agreed-on team-training curriculum.⁸

Baker et al., who recently evaluated several medical team-training programs, recommended that the health care industry establish standard, generic “teamwork related knowledge, skills, and attitude competencies,” leveraging all available research on teamwork and assessment tools.¹⁴ However, a standard, generic set of teamwork behaviors has yet to be established and an effective way to assess what has been taught remains controversial.

Research on teamwork behaviors in nonmedical industries has done much to inform the application of observation tools in numerous specialty-specific health care teams.^{10–12} In this article we describe our efforts to further this work by developing a behavior-based tool to quantitatively assess communication and team skills of health care providers in a variety of real and simulated clinical events.

Observing Team Performance: Its Evolution in Health Care

Observing team performance in health care is based on research in behavioral psychology and organizational behavior, human factors, high-reliability organizations, and patient safety. Studies of safety in the work place can be traced back to the military during the 1940s, ultimately diffusing to nonmilitary industries, especially during the 1990s.^{15,16} Evaluating team behaviors during routine and crisis situations has allowed for a better understanding of the effectiveness of industry-specific team-training interventions.¹⁷ Observing team behaviors in real time and evaluating how teammates interact with one another toward a common goal is fundamental to assessing the effectiveness

of interventions aimed at improving team skills.⁶

Now in its fifth generation, crisis resource management (CRM) is employed by crew members to identify existing and potential threats of error and to develop, communicate, and implement plans to avoid or mitigate these threats.^{18,19} CRM training also aims to improve morale and enhance efficiency of operations. Aviation and other non-health care industries supplement their technical training programs with CRM training. Proof that errors are interrupted or prevented as a direct result of CRM behaviors has yet to be established. However, observation of team performance conducted in a consistent, standardized fashion by skilled observers are methods by which high-hazard industries assess CRM behaviors in simulated settings—such as Line Operation Safety Audits (LOSA), a “no-jeopardy” observation of crew behaviors and situational factors—in real-life settings.²⁰ Similarly, a “non-technical skills assessment tool (NOTECHS) is used to assess the presence of behaviors that CRM training aims to embed in the culture of the flight deck.”²¹

Behavior-based line audits or “operational report cards” document the proficiency of team skills such as leadership, communication, and work-load management and monitor and cross-check performance. Conducting observations of teamwork requires a contextual appreciation for how these skills should be demonstrated during routine and critical events. This understanding then guides the development of behavior-based measures and the educational process by which observers learn to use these methods in a consistent, standardized fashion. The resulting behavior-based reports reflect the extent to which teamwork behaviors are exhibited in the working environment—and how they change following a CRM training program.

Observing and evaluating teamwork behavior in health care settings is relatively infrequent compared to other industries. Medical teams, like flight crews, generally assemble ad hoc. Just as a pilot scheduled to fly at a moment’s notice may be paired with a navigation officer who he or she has never met, so, too, may a surgeon perform emergency surgery in an unfamiliar operating room (OR) with unfamiliar personnel from the night shift. Much could be learned about medical errors and error threat and management by directly observing health care team performance. Of the various methods of observation

currently in use, most focus on some combination of technical or task-oriented skills and teamwork behaviors exhibited by established, “intact” teams and temporary, “ad hoc” teams or work groups.

For example, Morey et al. used their CRM-based, behaviorally anchored rating scale, the “Team Dimensions Rating Form,” for qualitative assessment of team behaviors exhibited by emergency medicine health care providers.²² Team behaviors, grouped in five team dimensions (maintain team structure and climate, plan and problem solve, communicate with team, manage work load, and improve team skills) were scored on a 7-point Likert scale. Specific team behaviors observed included establishing a team leader, assigning roles and responsibilities, communicating essential information, mutual accountability and respect, conflict resolution, using structured communication and critical terms or language, cross-monitoring, maintaining situational awareness, and team/shift reviews and updates.

Healey, Undre, and Vincent’s Observational Teamwork Assessment of Surgery (OTAS) tool²³ assessed “what (surgical teams) do and how they do it” by recording certain clinical, technical, and interpersonal skills and behaviors that occur when team members interact with the patient and one another. For 50 surgical procedures, a surgeon observed task-related elements (planning, availability of appropriate equipment, completion of appropriate checklists, and completion of communications relative to information about the patient) and a psychologist focused on teamwork behaviors within the five dimensions of cooperation, coordination, leadership, monitoring, and communication.

Thomas, Sexton, and Helmreich²⁴ developed a teamwork audit for neonatal resuscitation, drawing from line audits used to assess pilots’ CRM skills. Ten behavioral markers, including information sharing, intentions shared, managing work load, vigilance/environmental awareness, and overall teamwork and leadership, were identified from interviews, surveys of providers, and video observations of neonatal resuscitations.

Flin and Maran²⁵ developed the Anesthesiologist Non-Technical Skills (ANTS) tool, a behaviorally anchored teamwork assessment tool derived from attitudinal surveys of anesthesiologists, real-time observations of anesthesiologists caring for patients, and quality assurance reviews of critical incidents with adverse outcomes. ANTS divides nontechnical skills into four general categories—task man-

agement, team working, situation awareness, and decision making—each with 15 elements or behavioral markers. Observations of anesthesiologists in simulated OR scenarios informed the development of an educational program, called Crisis Avoidance and Resource Management for Anesthesiologists (CARMA), for better management of events. Through course evaluation and debriefing, the researchers found that the behavioral markers identified in ANTS could be used to assess the nontechnical skills of other medical providers with whom the anesthesiologist interacts.

We could not identify in the literature, as exemplified by these studies, a behavior-based assessment tool suitable for gauging teamwork skills that are not specialty-specific but are widely applicable and reflect good practice across the health care professions. The CATS Assessment was specifically developed for this purpose.

The CATS Assessment

The Communication and Teamwork Skills (CATS) Assessment (Figure 1, page 555), which was based on the work that we have briefly surveyed, was then designed, tested through several rapid improvement cycles, and piloted in various settings throughout the Partners HealthCare system. Our goal was to develop an instrument that provides evidence of the presence and quality of particular skills and also reinforces those skills by providing useful feedback to the observed teams.

BEHAVIOR MARKERS AND SCORING SYSTEM

Behavior Markers. The behavior markers in the CATS Assessment were selected from CRM behavior-based markers used in aviation and the military, which have been then adapted to health care.^{1,7–9,24} The LOSA, ANTS, and OTAS tools are closely aligned and share four team behavior domains—task management, teamworking (coordination and cooperation), situation awareness, and decision making. These four domains are subdivided into various elements consisting of planning and preparation, prioritization, execution, identifying and using resources, coordinating team activities, communicating and exchanging information, assertiveness and authority, assessing capabilities, supporting others, gathering information, understanding and recognition, anticipation, identifying options, balancing risks and selecting options, and re-eval-

Table 1. Comparison of LOSA Team Assessment Domains with Those of Behaviorally Anchored Health Care Team Assessment Tools*

Team Assessment Domains

LOSA	ANTS	OTAS	CATS
Briefings			Situational awareness
Plans stated	Task management		Communication
Work load assignment	Task management	Coordination	
Contingency management	Situation awareness	Leadership	Situational awareness
Monitor/cross-check	Teamworking	Awareness	Cooperation
Work load management	Task management	Coordination	
Vigilance	Situation awareness	Awareness	Situational awareness
Automation management	Task management		
Evaluation of plans	Decision making	Awareness	Coordination
Inquiry	Teamworking	Cooperation	Cooperation
Assertiveness	Teamworking	Leadership	Cooperation
Communication environment	Teamworking	Communication	Communication
Leadership	Teamworking	Leadership	Coordination

* LOSA, Line Operation Safety Audits; ANTS, Anesthesiologist Non-Technical Skills; OTAS, Observational Teamwork Assessment of Surgery; CATS, Communication and Teamwork Skills.

uation. These elements encompass specific observable behaviors, some of which are ascribed to more than one element. However, 21 of these specific observable behaviors were common to ANTS and OTAS and were deemed applicable to all health professions. We incorporated these 21 specific behaviors into CATS because they reflect CRM-based behaviors that are increasingly being taught in health care and that facilitate good teamwork regardless of discipline.

We then clustered these 21 behaviors into four categories: coordination, cooperation, situational awareness, and communication. Leadership behaviors, regarded by some as a separate category, are contained within these categories and include establishing an “event manager” during a crisis.

Table 1 compares the domains among CATS and other behaviorally anchored health care team assessment tools with those LOSA, the behavior-based tool created to assess aviation team performance. Table 2 (at right) illustrates how the behavior-based categories or domains of CATS, ANTS, and OTAS compare with one another.

The Leadership Issue. We consider leadership, a significant element of teamwork, at two levels: senior leadership and situational leadership. Senior leaders, through organi-

Table 2. Comparison of Team Behavior Domains: ANTS, OTAS, and CATS*

ANTS	OTAS	CATS
Situation awareness	Awareness	Situation awareness
Decision making	Coordination	Coordination
Teamworking	Communication	Communication
Task management	Cooperation	Cooperation
	Leadership	

* ANTS, Anesthesiologist Non-Technical Skills; OTAS, Observational Teamwork Assessment of Surgery; CATS, Communication and Teamwork Skills.

zational systems and structures, facilitate a culture of non-negotiable mutual respect and trust so that effective teamwork can flourish. CATS, which is designed to measure specific communication and team practices, does not address senior leadership. Leadership practices are found at the level of situational leadership, a matter of behavior, not rank. Members of a team should be able to cede authority to the member who has the informed perspective, relevant information, or expertise for best leading the event, especially during a crisis. Effective situational leadership is not

Table 3. Definitions of Behavior Markers for Communication and Teamwork Skills (CATS) Assessment*

Briefing: A conversation and two-way dialogue of concise and relevant information shared prior to a procedure or activity. Surgical “time-out” may be a briefing. Elements: Get the person’s attention; make eye contact; introduce yourself; use names; use SBAR; supply explicitly asked for information; talk about next steps; encourage ongoing monitoring and cross-monitoring.

Verbalize plan: Speak aloud the next steps for the procedure and/or care of the patient.

Verbalize expected time frames: Speak aloud time frames for particular interventions. “We’ll give this another two minutes and if there’s no change we’ll try X.”

Debriefing: A conversation and two-way dialogue of concise and relevant information shared after the procedure or activity is completed. Debriefing identifies what went well, what could have been done differently, and what was learned.

Establish event manager if crisis arises: Verbally identify who’s in charge if situation becomes a crisis; event manager does not participate in active interventions but maintains situational awareness and verbalizes plans, needs, and time frames.

Visually scan environment: Clinicians look up, look at one another, look at equipment, and look around the room.

Verbalize adjustments in plan as changes occur: Speak aloud new plans, changes in strategy or intervention, and new time lines as procedure progresses.

Request additional external resources if needed: Speak aloud, asking for help from outside the team—other clinicians, rooms, equipment, consults, etc.

Ask for help from team as needed: Team members speak aloud, asking for assistance from members of the team.

Verbally request team input: Ask aloud for team’s suggestions, opinions, comments, or ideas.

Cross-monitoring: Acknowledge concerns of others—watching team members, awareness of their actions, verbally stating concerns, sharing work load, verbally updating others in a manner less formal than briefing, responding to the concerns of team members.

Speak up, verbal assertion: If team members are uncomfortable or unclear, they speak aloud their concerns and state an alternative viewpoint or suggest an alternative course of action. Individuals are sufficiently persistent to clearly state their opinions. If team members perceive something as unsafe, they speak aloud to indicate that. If responses to expressed concerns are not satisfactory and unsafe situations continue, individuals escalate the concern by bringing in other clinicians.

Closed-loop communication: When a request is made of team members, someone specifically affirms aloud that they will complete the task and states aloud when the task has been completed.

SBAR: Use of specific structured communication that states the situation, background, assessment, and recommendation.

Critical language: Use of key phrases understood by all team members to mean “stop and listen, we have a potential problem.” Specific phrases may differ from one institution or work unit to another.

Verbal updates of situation: Think aloud—Team members verbally state their perceptions, actions, and plans as the procedure progresses.

Use team members’ names: Use team members’ names.

Communicate with patient: Team members speak to and respond to the patient

Use appropriate tone of voice: Team members use a tone of voice that is calm, professional, and not unnecessarily loud.

* SBAR, situation-background-assessment-recommendation.

static or unique to an individual. CATS measures leadership practices within the team through briefings, verbalizing expected time frames, debriefings, verbalizing adjustments in plans as changes occur, requesting team input, receptivity to assertion, and verbalizing updates. Should a crisis arise, CATS has a section designed to capture the specific leadership skill for effective team coordination, namely, that of establishing an event manager. CATS is intended to evaluate team and not individual performance, and a behavior is scored for the team regardless of which team member exhibits the behavior. We hypothesize that CATS can be modified to specifically isolate leadership or other behavior categories when specifically

applied to individual assessment of teamwork skills.

Testing of Scoring Method. Before piloting CATS, many plan-do-study-act cycles of rapid improvement were completed.²⁶ The members of the quality improvement team, composed of two physicians [A.F., R.G.], a nurse [A.K.], and a nonclinical patient safety/quality improvement specialist [L.M.], scored six videotaped simulations, each involving a different team of clinicians. Three of the simulations concerned a situational crisis involving an emergency cesarean section with acute maternal compromise. The remaining three were code-training scenarios. The quality improvement team also scored three live observations of gastric bypass surgeries in a community

Table 4. Pre-Observation Checklist for Communication and Teamwork Skills (CATS) Assessment*

Communication and Teamwork Skills Assessment Checklist

(Complete just prior to observation)

- | | |
|--|--------------------------|
| Review definitions of behavior markers. | <input type="checkbox"/> |
| Always expected in all settings: briefing and debriefing | <input type="checkbox"/> |
| Always expected post-training: SBAR, closed loop | <input type="checkbox"/> |
| If crisis arises, use second section of CATS. | <input type="checkbox"/> |
| Place one mark in the appropriate column each time behavior is observed or expected. | <input type="checkbox"/> |
| Ask team about any critical language they use. | <input type="checkbox"/> |
| Introduce self and observation, as appropriate for setting. | <input type="checkbox"/> |

* SBAR, situation-background-assessment-recommendation.

hospital. These rapid-cycle tests were conducted from January through mid-March 2006. The team refined and adjusted the scoring and observation processes during the rapid-improvement cycles. For example, small details in the placement of columns and the graphic layout of behaviors affected the ability to score behaviors in real time.

The pilot observations were then conducted in a six-week period in April and May 2006 in three hospital settings, as follows:

- Nine observations of gastric bypass surgery (community hospital A)
- Six observations of interdisciplinary rounds (academic medical center)
- Two scheduled cesarean sections (community hospital B)

For gastric bypass surgery, the entire procedure was observed, from initially meeting the patient on the day of surgery to when responsibility of caring for the patient was transferred to postanesthesia care unit staff. Observers stood in the OR, watching, listening, and taking notes, without interacting with staff.

Findings from the nine gastric bypass surgeries were collated and discussed with the surgeons who participated, the perioperative nursing directors, and the chair of anesthesia. This generated discussion about collaborative efforts to further improve team function.

Scoring Method. CATS scoring requires real-time attention for tracking specific behaviors and clusters of behaviors during a routine or critical event. The scoring sheet is designed to allow the observer to mark each time specific behaviors occur and to grade their quality. Three columns are provided for this: “Observed and Good,” “Variation in Quality” (meaning incomplete or of variable quality), and “Expected but not Observed.”

Observers score behaviors on the degree to which the behavior meets the definition, as provided in a glossary (Table 3, page 553). When two or more individuals simultaneously observed, a pre-observation checklist (Table 4, at left), as well as a review of the glossary, was found helpful in maintaining consistency. For example, a briefing would be considered “Observed and Good” if concise and relevant information was exchanged and the individuals made good eye contact, used individual names, and explicitly mentioned the importance of ongoing monitoring, cross-monitoring, or speaking up (for example, “If you have concerns or see something that doesn’t make sense, speak up and say so”). A briefing of relevant information that lacked the use of team-related internal monitoring would receive a score of “Variation in Quality.”

Similarly, for the closed-loop communication behavior, if the response to a request for “Lasix 5 mg” is “Lasix 5 mg” the score would be “Observed and Good.” A response of “Lasix” or a simple “OK” or “Uh-huh” would be scored as “Variable in Quality,” and no response would be “Expected but not Observed.” Finally, if a situation arose in which an observer noted that one of the team members was being momentarily overwhelmed by increased work load, the cross-monitoring behavior would be considered “Observed and Good” if another team member spoke directly to that individual and offered assistance. In a response rated “Variable in Quality,” a team member would, for example, note the person’s work load without offering assistance. No mention by any team members of the work load would be considered “Expected but not Observed.”

Three behaviors were identified as not consistently applicable in routine, noncrisis situations—“establishing an event manager,” “escalation of asserted concern,” and “critical language.” These behaviors were positioned at the bottom of the CATS tool for use during critical events or if a routine event became critical.

Communication and Teamwork Skills (CATS) Assessment Instrument, Initial Version

Communication and Teamwork Skills Assessment

Date: _____ Observer ID: _____ Observation Start Time: _____ Observation End Time: _____

Case ID: _____ Procedure: _____

Category	Behavior	Observed and Good	Variation in Quality	Expected But not Observed	Comments
Coordination	Briefing				
	Verbalize plan				
	Verbalize expected timeframes				
	Debriefing				
Situational Awareness	Visually scan environment				
	Verbalize adjustments in plan as changes occur				
Cooperation	Request external resources if needed				
	Ask for help from team as needed				
	Verbally request team input				
	Cross Monitoring				
	Verbal assertion				
	Receptive to assertion and ideas				
Communication	Closed loop				
	SBAR				
	Verbal updates – think aloud				
	Use names				
	Communicate with patient				
	Appropriate tone of voice				

If Crisis Situation Arises					
Category	Behavior	Observed and Good	Variation in Quality	Expected But not Observed	Comments
Coordination	Establish event manager				
Cooperation	Escalation of asserted concern				
Communication	Critical language				

* SBAR, situation-background-assessment-recommendation.

Figure 1. Each time a behavior is either observed (as good or variable) or expected but not observed, a mark goes into the appropriate column. Space is also available for free-text comments and field notes to clarify the observation and provide examples of the behaviors.

Overall Communication and Teamwork Skills (CATS) Score for Each of the Nine Observed Gastric Bypass Surgeries

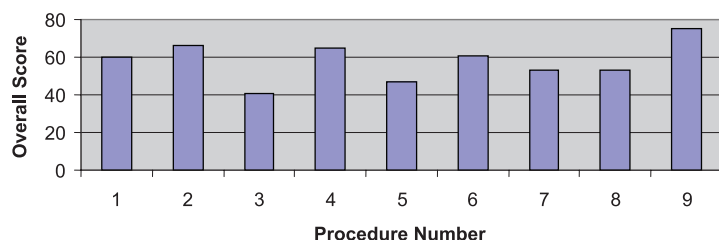


Figure 2. The overall Communication and Teamwork Skills (CATS) Assessment scores are shown for the nine observed gastric bypass surgeries. All the behaviors were averaged to obtain a combined score for each observation.

Refining the Scoring System. A method for weighting the observation scores was refined in consultation with a biostatistician and a certified behavior analyst. For each behavior, a weighted total is obtained as follows:

■ Marks in the “Observed and Good” column = 1.

■ Marks in the “Variation in Quality” column = .5.

■ Marks in the “Expected but not Observed” column = 0.

Scores are added together to achieve a weighted total. Thereafter, a second total is obtained by simply adding up *the total number of marks made*. The weighted-total, divided by total-number-of-marks, adjusted to a 100-point scale, is the quality score for that behavior. In this manner a quality score is established for each behavior during each observation period.

Graphic Display of the Scoring Process

Display of information is a key component to getting buy-in for improvement efforts and for giving feedback once improvement efforts are underway. Examples of graphical displays of baseline observations for nine individual gastric bypass surgeries (each surgical procedure lasting about two hours) are provided as follows:

■ For Each Observation Period:

–Overall average CATS scores (Figure 2, at left)

–Category scores (with overall average category score; Figure 3, page 556)

–A specific behavior quality score (Figure 4, page 557)

■ For All Observation Periods:

–Average Behavior Scores (Figure 5, page 556)

The final figure depicts a Behavior score over a time series with annotations, beginning after training was started (Figure 6, page 557)

The graphs begin with an overview of the nine individual gastric bypass surgeries observed, showing scores obtained by averaging all the behavior weighted-scores together (Figure 2). The figures then become progres-

Category Scores for Each Observation of Gastric Bypass Surgery

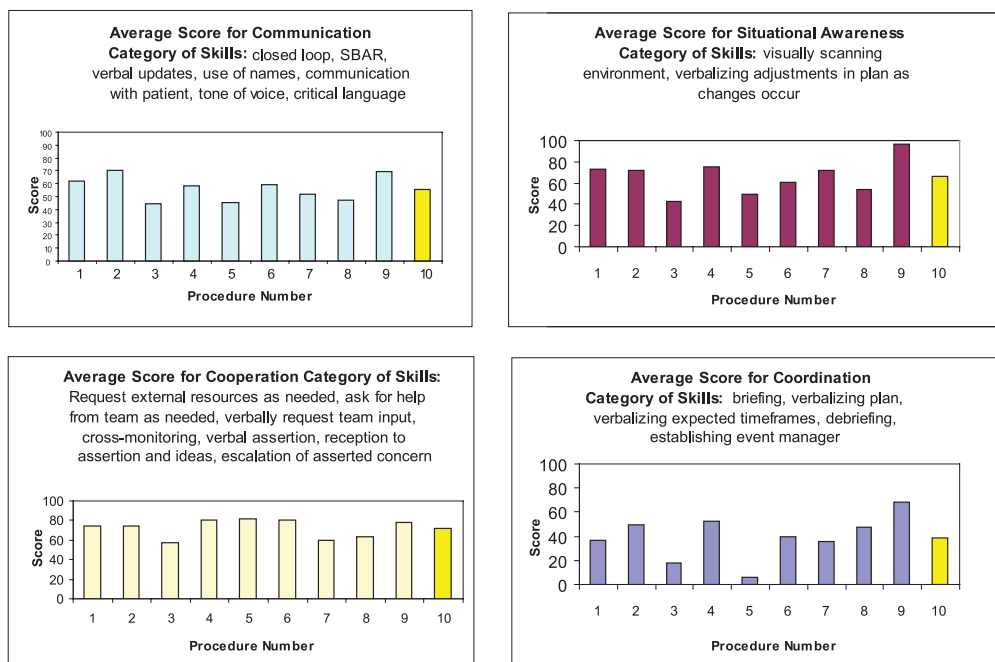


Figure 3. Behaviors comprising each category are listed. The bar on the far right (10) is an overall average of the nine observations. SBAR, situation-background-assessment-recommendation.

Example of Team Behavior Quality Score for Each Observation of Gastric Bypass Surgery

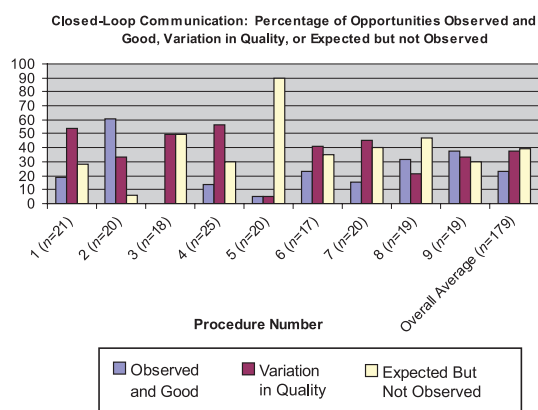


Figure 4. The scores for a single behavior—closed-loop communication—is shown for each procedure, based on whether it was judged “Observed and Good,” “Variation in Quality,” or “Expected But Not Observed.”

sively more detailed in stepwise fashion. The aggregate averages are of limited use in targeting areas for improvement but can function as an internal benchmark, allowing the group to compare overall scores over time.

Comparisons of scores at the category level are shown in Figure 3. Each category score represents the average of scores for all behaviors in that category. For example, the coordination category includes four behaviors: briefing, verbalizing plan, verbalizing expected time frames, and debriefing. The scores for each of the four behaviors were averaged to achieve the coordination category score. As can be seen in the example displayed in Figure 3, scores for the category of coordination range from a low (6) in procedure 5 to a high (68) in procedure 8. The average score for the coordination category for all observations combined is 39.

The quality of one specific behavior (closed-loop communication) as observed in each procedure is shown in Figure 4. For each of the nine procedures observed, a score was determined for this behavior based upon the percent-

Average Score for Each Behavior Across the Nine Observations of Gastric Bypass Surgery

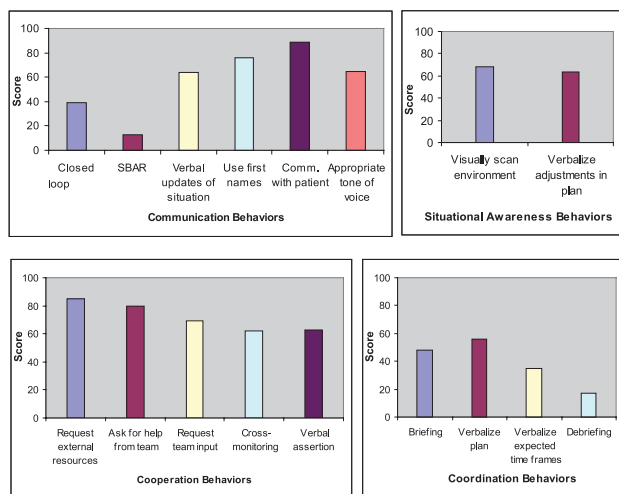


Figure 5. Each bar represents the average score for that behavior during all nine observations.

Hypothetical Statistical Process Control Chart for Briefings

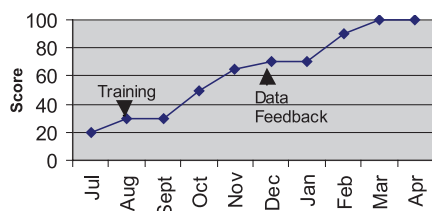


Figure 6. Change in performance over time on the briefings behavior is shown in this hypothetical statistical process control chart, with annotations.

age of opportunities for closed-loop communication. Throughout the procedure, the observer noted the number of appropriate opportunities for closed-loop communication (n) and scored each opportunity as either observed and good, variable in quality, or expected but not observed. In this example, in procedure 1, 21 opportunities for closed-loop communication were observed—of which 18.5% were either observed and good, 53.5% were

variable in quality, and 28% were expected but not observed. A similar analysis can be provided for each behavior included in the CATS assessment.

Behavior marker scores are shown in Figure 5 (at left). Finally, Figure 6 (at left) depicts a statistical process control chart, indicating changes over time in a specific behavior (briefings), with annotations indicating when improvement efforts were implemented.

Using the Scores. The CATS Assessment results enable clinicians to view a spectrum of scores—from overall score to the category-specific behaviors. We expect that feedback from a series of observations will assist teams in identifying and targeting behaviors for improvement and training. These data are amenable to statistical process control charting, providing motivation for improvement and reinforcement of desired behaviors.

An example of the summary opening paragraph of a feedback report follows, which would be attached to a series of graphs similar to those described earlier: “We conducted eight observations on the unit during November and December 2006, some of general activity and some of rounds and report. Overall scores for teamwork and communication behaviors were rather high, with significant variation between observations. Overall scores ranged between 40 and 77. Of the two lowest scoring observation events, one was of general activity and another of rounds/report. . . . Lowest scores by behavior, indicating potential opportunities for improvement, include debriefing, use of SBAR, and closed-loop communication. Highest scoring behaviors, indicating possible best practices, include using an appropriate tone of voice, cross-monitoring, use of team members’ names, and communicating appropriately with patients.”

Conclusion

In response to findings that suggest that health care providers do not function well as teams or communicate in a manner that minimizes threats and errors, health care organizations are increasingly providing training in effective communication and teamwork. Some curricula are based on CRM, others on conflict resolution and negotiation. Yet the “best” components of training are unclear, the organizational structure to ensure sustainability has not been defined, and the relationship between communication and teamwork behaviors and patient outcomes

remains unknown. A standard measure for assessing team performance must be applicable to a wide variety of health care settings. The CATS Assessment tool requires statistical validation and further study in its use in quantifying team behaviors and identifying opportunities for improvement. We invite researchers, clinicians, and educators involved in team training to use observation measures to better evaluate their training programs and continue to research and improve observation methodology. Providing quantifiable, objective feedback to their clinicians and organizations could facilitate optimal medical error prevention and management and promote a culture of safety. **J**

The authors thank Stuart Lipsitz, Ph.D., and Felix Santiago for their assistance in developing useful scoring mechanisms; the Center for Medical Simulation, Cambridge, Massachusetts, and STRATUS Center for Medical Simulation, Boston, which kindly allowed us to use videotaped simulations in the initial development of our observation methods.

Allan Frankel, M.D., is Director of Patient Safety, Partners HealthCare, Boston, and Senior Faculty, Institute for Healthcare Improvement, Cambridge, Massachusetts. **Roxane Gardner, M.D., M.P.H.**, is Assistant Professor of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, and Simulation Specialist, Center for Medical Simulation, Cambridge. **Laura Maynard, M.Div.**, is Administrative Director, Patient Safety Center, Duke University Health System, Durham, North Carolina. **Andrea Kelly, R.N., J.D.**, is Patient Safety Program Manager, Partners HealthCare. Please address requests for reprints to Allan Frankel, AFRANKEL@PARTNERS.ORG.

References

- Leonard M., Graham S., Bonacum D.: The human factor: The critical importance of effective teamwork and communication in providing safe care. *Qual Saf Health Care* 13:i85-i90, Oct. 2004.
- Thomas E.J., Sexton B.J., Helmreich R.L.: Discrepant attitudes about teamwork among critical care nurses and physicians. *Crit Care Med* 31: 956-959, Mar. 2003.
- Thomas E.J., Sherwood G.D., Helmreich R.L.: Lessons from aviation: Teamwork to improve patient safety. *Nurs Econ* 21:241-243, Sep.-Oct. 2003.
- Groff H., Martin P.B. (eds.): Understanding CRICO's perinatal claims. *Forum* 21:1-14, Mar. 2001.
- The Joint Commission: Preventing infant death and injury during delivery. *Sentinel Event Alert* Jul. 21, 2004. http://www.jointcommission.org/SentinelEvents/SentinelEventAlert/sea_30.htm (last accessed Jul. 11, 2007).
- The Joint Commission: *Sentinel Event Statistics-Dec. 31, 2006*. <http://www.jointcommission.org/SentinelEvents/Statistics/> (last accessed Jul. 11, 2007).
- White A.A., et al.: Cause and effect analysis of closed claims in obstetrics and gynecology. *Obstet Gynecol* 105:1031-1038, May 2005.
- Kohn L.T., Corrigan J.M., Donaldson M.S. (eds.): *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy Press, 2000.
- Baker D.P., et al.: Medical Team Training Programs in Health Care. *Advances in Patient Safety* vol. 4, Agency for Healthcare Research and Quality, Feb. 2005. <http://www.ahrq.gov/downloads/pub/advances/vol4/Baker.pdf> (accessed Jul. 12, 2007).
- Lingard L., et al.: Communication failures in the operating room: An observational classification of recurrent types and effects. *Qual Saf Health Care* 13:330-334, Oct. 2004.
- Carthey J.: The role of structured observational research in health care. *Qual Saf Health Care* 12(suppl. 2):ii13-ii16, Dec. 2003.
- Barker K., Flynn E.A., Pepper G.: Observation method of detecting medication errors. *Am J Health Syst Pharm* 59:2314-2316, Dec. 2002.
- Salas E., Bowers C.A., Cannon-Bowers J.A.: Military team research: 10 years of progress. *Military Psychology* 7(2):55-75, 1995.
- Baker D.P., et al.: Literature Review: *Medical Teamwork and Patient Safety: The Evidence-Based Relation*. Publication No. 05-0053, Agency for Healthcare Research and Quality, Apr. 2005. <http://www.ahrq.gov/qual/medteam/> (accessed Jul. 12, 2007).
- Cooper G.E., White M.D., Lauber J.K.: *Resource management on the flightdeck: Proceedings of a NASA/Industry Workshop*. NASA Conference Publication No. CP-2120. Moffett Field, CA: NASA-Ames Research Center, 1980.
- Lauber J.K.: Cockpit resource management: Background and overview. In Orlady H.W., Foushee H.C. (eds.): *Cockpit Resource Management Training: Proceedings of the NASA/MAC Workshop*. NASA Conference Publication No. 2455. Moffett Field, CA: NASA-Ames Research Center, 1987.
- Helmreich R.L., Wilhelm J.A.: Outcomes of crew resource management training. *International Journal of Aviation Psychology* 1(4):287-300, 1991.
- Helmreich R.L., Merritt A.C., Wilhelm J.A.: The evolution of Crew Resource Management training in commercial aviation. *International Journal of Aviation Psychology* 9(1):19-32, 1999.
- Wiener E.L., Kanki B.G., Helmreich R.L.: *Cockpit Resource Management*. San Diego: Academic Press, 1993.
- Klinect J.R., et al.: Line Operations Safety Audit (LOSA): Definition and operating characteristics. In Proceedings of the 12th International Symposium on Aviation Psychology. Dayton, OH: Ohio State University, 2003, pp. 663-668.
- Avermaete J.A.G. van.: NOTECHS: Non-Technical Skill Evaluation in JAR-FCL; Nationaal Lucht-en Ruimtevaartlaboratorium, National Aerospace Laboratory NLR-TP-98518, 1988. <http://www.nlr.nl/smartsite.dws?id=4646> (accessed Jul. 2007).
- Morey J.C., et al.: Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the MedTeams project. *Health Services Research* 37:1553-1581, Dec. 2002.
- Healey A.N., Undre S., Vincent C.A.: Developing observational measures of performance in surgical teams. *Qual Saf Health Care* 13:i33-i40, Oct. 2004.
- Thomas E.J., Sexton J.B., Helmreich R.L.: Translating teamwork behaviours from aviation to healthcare: Development of behavioural markers for neonatal resuscitation. *Qual Saf Health Care* 13:i57-i64, Oct. 2004.
- Flin R., Maran N.: Identifying and training non-technical skills for teams in acute medicine. *Qual Saf Health Care* 13:i80-i84, Oct. 2004.
- Langley G., et al.: *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*. San Francisco: Jossey-Bass, 1996.