Defining Team Performance for Simulation-based Training: Methodology, Metrics, and Opportunities for Emergency Medicine

Marc J. Shapiro, MD, Roxane Gardner, MD, MPH, Steven A Godwin, MD, Gregory D. Jay, MD, PhD, David G. Lindquist, MD, Mary L. Salisbury, RN, MSN, Eduardo Salas, PhD

Abstract

Across health care, teamwork is a critical element for effective patient care. Yet, numerous well-intentioned training programs may fail to achieve the desired outcomes in team performance. Hope for the improvement of teamwork in health care is provided by the success of the aviation and military communities in utilizing simulation-based training (SBT) for training and evaluating teams. This consensus paper 1) proposes a scientifically based methodology for SBT design and evaluation, 2) reviews existing team performance metrics in health care along with recommendations, and 3) focuses on leadership as a target for SBT because it has a high likelihood to improve many team processes and ultimately performance. It is hoped that this discussion will assist those in emergency medicine (EM) and the larger health care field in the design and delivery of SBT for training and evaluating teamwork.


Keywords: team, training, simulation, emergency medicine, metrics, leadership

Health care delivery is an extremely complex operation, and despite the best intentions of many dedicated professionals, we are constantly reminded of the perils of the significant errors in reported sentinel events. Newsworthy errors, such as wrong-site surgery, highlight deficiencies in nontechnical skills and behaviors including teamwork. Across health care, teamwork is a critical element for effective patient care. Yet, reports reviewing the effectiveness of healthcare teams have consistently supported the saying “a team of experts does not necessarily constitute an expert team.” The truth to this statement is demonstrated by a 2004 report indicating that 72% of the root-cause analyses of sentinel events in perinatal units across the United States identified failures in communication (a core teamwork skill) as a contributor. Similarly compelling data were obtained by a needs analysis conducted during the MedTeams™ project. Approximately 40% of emergency department (ED) closed claim cases involve teamwork errors that could have been mitigated or prevented.

In response to such statistics, many health care organizations have attempted to implement team training programs. Some have succeeded, while others have struggled to achieve desired outcomes. Hope for the improvement of teamwork in health care is provided by the success of the aviation and military communities in utilizing simulation-based training (SBT) for training and evaluating teams.

The earliest attempt to codify teamwork behaviors and their measurement among emergency care providers occurred during the MedTeams™ project, a multicenter prospective investigation of crew resource management (CRM) from 1996 to 2001. The results from the MedTeams™ project showed that an increase in safety-related behaviors is achievable through a well-constructed didactic curriculum designed to inculcate team behaviors in groups of clinicians. Teamwork was measured through 1) behavioral anchored rating scales...
Simulation-based Team Training

**Principle 1: Establish the Team-based Competencies to Be Assessed and Trained—The Learning Outcomes.** The science of training instructs us that the first step of any training program is to identify what needs to be trained. Training for teamwork in healthcare is no different. The Accreditation Council for Graduate Medical Education (ACGME) requires that residents develop proficiency in six core competencies (e.g., KSAs). Although “teamwork” is not overtly specified as one of these competencies, elements of teamwork are incorporated (e.g., the development of effective communication skills is the primary emphasis under the “interpersonal and communication skills” competency). Thus, teamwork relevant aspects of these competencies should be considered in SBT design.

**Figure 1.** Overview of the Simulation Module for Assessment of Resident Targeted Event Responses (SMARTER) processes for training teamwork in EM. Adapted from Rosen et al.®

EM = emergency medicine; KSAs = knowledge, skills, and attitudes.
In addition to the ACGME, other sources highlight the teamwork competencies most critical and commonly required within EM. Over 20 years of theoretical and empirical literature on team training abound. Emphasized across this literature are three key components for health care teamwork: communication, coordination, and cooperation. To obtain a more detailed, systematic view of teamwork competencies, team task analysis identifies the components of a task that are related to task work versus teamwork. Ultimately, tasks related to teamwork can be used to estimate the frequency and relative importance of team member coordination during specific tasks. Input from those considered “experts” in their position is often gathered as part of this process. SBT for teamwork will be most beneficial when scenarios are designed to target the team-based KSAs identified as common and critical for performance. The competencies selected for training serve as the content for the design of SBT scenarios (see Principle 2) and performance measurement tools (see Principle 3).

**Principle 2: Provide Opportunities for Guided Teamwork Practice—Carefully Craft Scenarios.** Knowing about teamwork does not mean one will behave as a team member. Health care providers must be able to perform teamwork skills. SBT provides a setting for the hands-on guided practice of teamwork behaviors to help ensure skill acquisition. Continued guided practice leads to learning; however, not all practice is good practice. Mere repetition does not guarantee the accurate development of teamwork skills. Rather, practice must be guided to ensure that teamwork skills are accurately developed. Practice opportunities should focus on a specific skill, with immediate provision of feedback about what areas can be improved and how these improvements can be made.

In conducting SBT, the ability to provide structured and focused practice opportunities begins during scenario design. An effective SBT scenario targets only a small subset of the desired team training objectives. To ensure that trainees practice and demonstrate the specific KSAs of interest, a scenario must contain “triggers” designed to elicit a given teamwork response. These triggers prompt team members to demonstrate a specific behavior. For example, in a team scenario designed to elicit communication, a team member may be instructed by a confederate to administer a medication to a patient for which the team was previously told the patient was allergic. The team member would be expected to voice concern about the medication and remind the confederate about the allergy. Any action other than this by the team member would potentially result in a decline in the patient’s condition. Following the simulation, immediate feedback, including specific positive and negative examples, should be provided.

**Principle 3: Measure Team Performance—Diagnose Teamwork Strengths and Weaknesses.** Measurement creates an opportunity for learning. SBT works when measurement problems (e.g., observation protocols, behavioral markers) are designed to diagnose team performance. Metrics, behavioral markers, and observation protocols are a must for effective SBT. By designing diagnostic measures that directly parallel the trigger events and targeted teamwork concepts embedded within the SBT teamwork simulations, observers know exactly what they are watching for and when a behavior should occur (see Figure 2). The subjectivity of performance is therefore minimized. Furthermore, designing measurement tools assists in ensuring thorough and relevant feedback, because a tool provides a structured template. Debriefing can focus specifically on whether expected behaviors did or did not occur.

**Principle 4: Develop Robust Debriefing Protocols—Link Feedback to Learning Outcomes.** Feedback is, of course, essential in the team learning process. Feedback informs the team about how they are performing and identifies strengths and weaknesses. To be most effective, feedback must be accurate, specific, developmental, and provided immediately following scenario participation in a structured debriefing session. A facilitator (typically the trainer) should use a diagnostic scenario-based tool to direct debriefings; however, all team members should actively participate and contribute to the identification of learning opportunities and suggestions for remediation. For open and honest participation to occur, debrief sessions must be perceived as safe environments where learning is the shared objective, rather than criticism or placing blame. Team members should discuss aspects of performance in the scenario that went well and the factors that contributed to the desired outcomes. This allows the team to focus on repeating these successful behaviors in future performance sessions. Remediation opportunities may involve the repetition of a previously performed scenario or a new scenario requiring a similar set of teamwork skills.

Teamwork is too complex to be taught in a single scenario. Rather, a collection of scenarios should be provided to help ensure skill acquisition. SBT works when immediate feedback allows trainees to learn from their mistakes and receive suggestions for improvement.

---

**Figure 2.** Example of an observational checklist for the team competency of “leadership.” “Hits” are a dichotomous scoring of whether or not the targeted behavior was observed. IG = instructor guided; used when the behavior is observed but was coached by a trainer. Adapted from Rosen et al. ED = emergency department; RN = registered nurse.
created, with each scenario only focusing on a small, specific subset of the teamwork KSAs. Maintaining a library of scenarios is also beneficial in that team training is not a “one size fits all” process. No two teams are identical, and individual teams will change over time as new technologies are introduced and medical procedures and knowledge evolve. Scenarios emphasizing a limited set of teamwork competencies allow SBT facilitators to better accommodate and focus on the unique training needs of a given team.

TEAM PERFORMANCE METRICS IN HEALTH CARE: REVIEW OF THE LITERATURE AND RECOMMENDATIONS

What Are Team Performance Metrics and Why Use Them?
The most common reason for evaluation is to determine the effectiveness or impact of an intervention and the ways in which it can be improved or strengthened. However, the information obtained from evaluation is only that, information. To yield insight, the quality of information must be accurate, reliably high, and obtained through the use of high-performing instruments.

For the purposes of this discussion, the team is identified as a co-located group of two or more individuals working to plan, problem solve, and carry out safe, quality care across a population of patients, over time, within the context of high-stress, high-stakes, time-compressed environments where information is unavailable or uncertain, such as in the ED; or within the confines of a structured simulated setting. Teamwork is a complex construct requiring metrics broad enough to capture the cycles of performance, yet sensitive and precise enough to assess a single behavior.

No single measure can capture overall individual or team performance, and it is helpful to measure both team processes and outcomes to avoid the pitfalls in relying on “one single inadequate criterion.”

Thus, it is helpful to begin with the end in mind. The noise of team complexity is then reduced, precision gained, and the yield improved when evaluation is event-driven. In simulation, events are determined by the researcher to yield specific outcomes. Scenarios are scripted, and performance criteria are determined with the intent to run sequential scenarios of increasing difficulty. Instruments commonly composing team performance evaluation are perception surveys, team performance observations, combined with process and outcome measures. Which instruments to choose is a matter of science.

Brief History of Team Performance Evaluation in Health Care
The value of conducting direct, real-time observations as a meaningful way to assess the caliber of teamwork behaviors in health care settings is becoming widely appreciated. However, this methodology is still in its infancy. Gauging performance through direct observation of team behaviors is grounded in the knowledge gained through research in behavioral psychology, organizational behavior, human factors, and industrial and occupational safety. Since the 1940s, the military has studied safety in the work place, leading to similar safety-driven initiatives by nonmilitary, high-risk industries.

Lauber first defined the term “cockpit (crew) resource management” (CRM) to mean “using all available sources—information, equipment, and people—to achieve safe and efficient flight operations.” CRM has expanded to include the identification of potential threats of error, to avoid or mitigate these threats, and to improve morale and enhance efficiency of operations. Aviation and other non–health care industries supplement their technical training programs with CRM training. Behavior-based aviation safety audits, such as line-oriented flight training and line operation safety audits (LOSA) document team skill proficiencies of leadership, communication, workload management, and monitor/cross-check performance. A “nontechnical skills” assessment tool was developed specifically to evaluate CRM behaviors of individual pilots during flight deck maneuvers. There is tension between observing teams versus an individual’s performance, because an individual’s performance can affect the team’s performance. Metrics used to assess teams may not be appropriate for assessing an individual teammate, and vice versa.

Within the health care arena, observing teamwork while managing routine or critical events can reveal much about medical errors and identify ways to prevent or mitigate their effects. A contextual understanding for how teamwork skills should be demonstrated during routine and critical events is needed to guide the development of standardized, behavior-based metrics. Gaba et al. first adapted aviation-based CRM to anesthesia in 1989 and named it “anesthesia crisis resource management.” Flin and Martin identified five key elements to consider when developing behavioral-based assessment tools for CRM skills applicable to aviation and anesthesia teams. These elements include the unit of assessment, identification of team skills and their markers, the method of assessment, rater reliability, and rater training. Translating aviation-based CRM principles for use in anesthesia triggered the adoption and spread of these techniques and use of behavior-based assessments across the medical and surgical specialties. Morey et al. developed and used a CRM-based, behaviorally anchored rating scale known as the “team dimensions rating form” for qualitatively assessing team behaviors of EM providers. The team dimensions form grouped team behaviors in five dimensions (maintain team structure and climate, plan and problem solve, communicate with team, manage workload, improve team skills) and scored them on a 7-point Likert scale for specific team behaviors, such as team leadership, role clarity, structured communication, mutual accountability and respect, conflict resolution, cross-monitoring, situational awareness, and team/shift updates.

Healey et al. created an “observational teamwork assessment of surgery” (OTAS) tool to assess “what (surgical teams) do and how they do it.” They recorded specific clinical, technical, and interpersonal skills and team behaviors among themselves and with their
patient. Teams were assessed for task-related elements (planning, availability of appropriate equipment, completion of surgical checklists, and patient-related communication) and teamwork behaviors within the five dimensions of cooperation, coordination, leadership, monitoring, and communication.

Thomas et al.\(^3\) developed a teamwork audit for neonatal resuscitation, based on line audits used to assess pilots’ CRM skills. Ten behavioral markers, including information sharing, intentions shared, managing workload, vigilance/environmental awareness, and overall teamwork and leadership, were identified from interviews, surveys of providers, and video observations of neonatal resuscitations.

Flin and Maran\(^4\) 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 non-technical skills into four general categories of task management, team-working, situation awareness, and decision-making, each with 15 elements or behavioral markers.

**Current Use of Health Care–related Team Performance Metrics**

**Metrics Assessing Generic Team Work.** Most behavior-based teamwork assessment tools are specialty-specific, but Baker et al.\(^4\) recommended that all available resources be directed toward establishing standard, generic teamwork skills, and competencies for the health care industry. Frankel et al.\(^5\) developed a non specialty-specific behavior-based assessment tool known as the “communication and team skills” (CATS) assessment. This tool, based on CRM behaviors grounded in the aviation and military industries, was designed specifically for the purpose of gauging teamwork skills across health care professions.\(^31,32,36,39\) Currently undergoing statistical validation, the tool identifies whether expected team behaviors are present and evaluates their quality. The health care team assessment domains of CATS compares favorably with that of LOSA, OTAS, and ANTS (Table 1).\(^42\)

**Metrics Assessing Specialty-specific Team Work.** There is no one size fits all, criterion standard team performance metric in health care, but much research is being directed toward identifying those that are theoretically relevant and scientifically sound, capture competencies, focus on observable behaviors, and provide meaningful information that facilitates improved performance and outcomes.\(^15\) All forms of metrics are in use across the medical and surgical specialties, ranging from event-based, automated performance monitoring and self-report measures, to BARS and behavioral observation scales (BOS; Table 2).\(^15,24\)

There is also no criterion standard approach for the process by which specialty-specific team performance metrics are developed. However, a stepwise approach provides a suitable example of how such tools should be developed (Table 3). In developing surgical team performance measures, Yule et al.\(^43\) reviewed the literature for tools assessing nontechnical skills of surgeons in the operating room. They identified core categories of nontechnical skills that were measured, including communication, teamwork, leadership, and decision-making, but deemed these metrics to be psychometrically deficient. Flin et al.\(^44\) used attitudinal surveys to assess surgeon’s perception of safety in the operating room. The adverse surgical events were analyzed via critical incident interviews and direct observation of surgeons, which resulted in a taxonomy of teamwork behaviors and a behavior-based rating system to assess such skills.\(^45\) Reliability testing of their instrument is currently under way or nearing completion. Flin et al.\(^46\) then developed a nontechnical skills curriculum, to train surgeons and maximize safety and quality in the operating room. This stepwise process, although tedious and labor-intensive, provides a comprehensive and focused understanding of the demands of the

<table>
<thead>
<tr>
<th>Briefings</th>
<th>OTAS</th>
<th>ANTS</th>
<th>CATS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plans stated</td>
<td>Coordination</td>
<td>Task management</td>
<td>Situation awareness</td>
</tr>
<tr>
<td>Workload assignment</td>
<td>Leadership</td>
<td>Task management</td>
<td>Communication</td>
</tr>
<tr>
<td>Contingency management</td>
<td>Awareness</td>
<td>Situational awareness</td>
<td>Situation awareness</td>
</tr>
<tr>
<td>Monitor/crosscheck</td>
<td>Coordination</td>
<td>Team working</td>
<td>Cooperation</td>
</tr>
<tr>
<td>Workload management</td>
<td>Awareness</td>
<td>Task management</td>
<td>Communication</td>
</tr>
<tr>
<td>Vigilance</td>
<td>Awareness</td>
<td>Situational awareness</td>
<td>Task management</td>
</tr>
<tr>
<td>Automation management</td>
<td>Situation awareness</td>
<td>Coordination</td>
<td>Decision-making</td>
</tr>
<tr>
<td>Evaluation of plans</td>
<td>Awareness</td>
<td>Team working</td>
<td>Team working</td>
</tr>
<tr>
<td>Inquiry</td>
<td>Leadership</td>
<td>Team working</td>
<td>Team working</td>
</tr>
<tr>
<td>Assertiveness</td>
<td>Communication</td>
<td>Team working</td>
<td>Communication</td>
</tr>
<tr>
<td>Communication environment</td>
<td>Leadership</td>
<td>Team working</td>
<td>Coordination</td>
</tr>
<tr>
<td>Leadership</td>
<td>Communication</td>
<td>Team working</td>
<td>Communication</td>
</tr>
</tbody>
</table>

ANTS = anesthesiologist non-technical skills; CATS = communication and team skills; LOSA = line operation safety audits; OTAS = observational teamwork assessment of surgery.
specialty. Contextually relevant, behavior-based evaluation tools can help identify opportunities to improve team performance and determine if such educational interventions are successful.

How Well Do These Metrics Assess Team Performance?

There is no criterion standard team performance metric or set of metrics, just as there is no criterion standard team training curriculum across the health care disciplines. However, since “what gets measured gets done,” instruments that accurately and reliably characterize team performance are critical to facilitating excellence in teamwork behaviors that ultimately mitigates medical error and prevents adverse clinical events.

How well metrics currently in use assess team performance depends on: 1) what is being measured and under what circumstances—during real clinical events or simulated scenarios; 2) instrument design and how well the tool captures the behavior in question; 3) the frequency of measurement; 4) the evaluator’s expertise in team performance and their understanding and familiarity with the instrument; and 5) the interpretation of the information gathered and the feedback provided to team members.

It is doubtful that one metric is sufficient to adequately capture the complex nature of team performance. The work of Morey et al. characterizes an approach to assessing team performance via a panoply of process and outcome measures, a virtual scatter gun approach of measuring a variety of specific, mutually agreeable events sequentially, over time, and tracking the results to monitor change in behavior. They demonstrate improved team performance and a decrease in clinical errors. The combined works of Yule and Flin et al. (Table 3) also underscore the importance of a solid methodology and sound psychometrics in developing

<table>
<thead>
<tr>
<th>Method</th>
<th>Description</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event-based measurements</td>
<td>General method involving behavioral checklist that are linked to scenario events and specific KSAs being trained</td>
<td>Explicitly links measurement opportunities (i.e., scenario events), acceptable behaviors, and KSAs being trained</td>
<td>Development of measures can be time consuming compared to other approaches; and measurement tools must be developed for each scenario Observers may focus on specific, isolated behaviors and miss the big picture of how the team is performing</td>
</tr>
<tr>
<td>BARS</td>
<td>Provides brief description of behaviors as anchors associated with each particular rating</td>
<td>Focuses observer’s attention on predefined events Reduces latitude of judgment by the observer by focusing on specific observable behaviors Can be modified Facilitates accurate ratings by providing concrete examples of the behaviors</td>
<td></td>
</tr>
<tr>
<td>BOS</td>
<td>Generally uses a Likert-type scale to rate the frequency of occurrence in certain team processes</td>
<td>Avoids potential problem of BARS by focusing on typical performance</td>
<td>Requires rates to estimate frequencies, potentially influencing final ratings by recency and primacy of events</td>
</tr>
<tr>
<td>Self-report measures</td>
<td>Questionnaires administered to individuals</td>
<td>Well-suited to capture affective factors that influence team performance</td>
<td>Does not capture dynamic performance; translating individual scores to team level scores can be problematic</td>
</tr>
</tbody>
</table>

*Modified from Rosen et al.*

BOS = behavioral observation scales; BARS = behaviorally anchored rating scales; KSA = knowledge, skill, attitudes.

---

specialty. Contextually relevant, behavior-based evaluation tools can help identify opportunities to improve team performance and determine if such educational interventions are successful.

How Well Do These Metrics Assess Team Performance?

There is no criterion standard team performance metric or set of metrics, just as there is no criterion standard team training curriculum across the health care disciplines. However, since “what gets measured gets done,” instruments that accurately and reliably characterize team performance are critical to facilitating excellence in teamwork behaviors that ultimately mitigates medical error and prevents adverse clinical events.

How well metrics currently in use assess team performance depends on: 1) what is being measured and under what circumstances—during real clinical events or simulated scenarios; 2) instrument design and how well the tool captures the behavior in question; 3) the frequency of measurement; 4) the evaluator’s expertise in team performance and their understanding and familiarity with the instrument; and 5) the interpretation of the information gathered and the feedback provided to team members.

It is doubtful that one metric is sufficient to adequately capture the complex nature of team performance. The work of Morey et al. characterizes an approach to assessing team performance via a panoply of process and outcome measures, a virtual scatter gun approach of measuring a variety of specific, mutually agreeable events sequentially, over time, and tracking the results to monitor change in behavior. They demonstrate improved team performance and a decrease in clinical errors. The combined works of Yule and Flin et al. (Table 3) also underscore the importance of a solid methodology and sound psychometrics in developing

<table>
<thead>
<tr>
<th>Stepwise process</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review literature for tools assessing specialty-specific behavioral skills</td>
<td>Yule et al.*</td>
</tr>
<tr>
<td>Identify core categories measured—communication, teamwork, leadership, and decision-making</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Survey attitudes and perceptions of safety</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Develop a taxonomy of specialty-specific teamwork behaviors</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Analyzing adverse specialty-specific events</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Conducting critical incident reviews</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Develop a behavior-based rating system</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Analyzing adverse specialty-specific events</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Conducting critical incident reviews</td>
<td>Flin et al.*</td>
</tr>
<tr>
<td>Develop a behavior-based skills curriculum</td>
<td>Flin et al.*</td>
</tr>
</tbody>
</table>
tools for assessing specialty-specific team performance. Well-contrived metrics can provide a robust reflection of teamwork behaviors and reveal opportunities to promote best practices for team performance that lead to reduction in medical errors and safer patient care.

Recommendations for Using Team Performance Metrics
The most common reason for evaluation is to determine first the effectiveness or impact of an intervention and next the ways in which the attribute or impact can be improved. The accuracy of determining effectiveness hinges on engaging a solid evaluation methodology and the use of high-performing instruments designed to precisely assess for the attribute sought.

The majority seeking to determine the impact of a teamwork intervention are not individuals engaged in research. Rather, they are individuals seeking to ensure the communication of knowledge, skills, and actions essential to high-functioning reliably safe organizations. To that end, there is a need for researcher and nonresearcher to partner and develop high performing measures, capable of assessing the attribute sought, while imposing little or no undue burden upon a complex dynamic environment.

SBT FOR LEADERSHIP
Strong and effective team leadership is critical to team performance as leaders assist in team development, problem-solving, and providing motivation. Just as teamwork consists of behaviors that can be taught, reinforced, and measured, team leadership may also be defined by such behaviors. Different situations require different leadership strengths, but there are global leadership skills and abilities that are transferable. In the medical setting, a good leader must not only be a good clinician, but also have a strong sense of team dynamics and understand the impact of personal interactions and communications.

The impact of leaders on team performance has been described by multiple sources. Marks et al. states that leaders integrate pertinent environmental cues into their existing cognitive structures to guide problem-solving. This then culminates into a course of action that serves to provide direction toward a given goal. Leadership behavior in teams may be subdivided into task-focused leadership, person-focused leadership, and task interdependence. Person-focused leadership addresses behavioral interactions, cognitive structures, and attitudes required for effective team interactions to facilitate successful task performance. Task interdependence defines the degree that team members depend on each other to perform a desired task. As team interdependence increases, team member performance must be coupled more intricately and increases the complexity of leadership. In 2000, Burke et al. conducted a meta-analysis, which suggested that the use of task-focused behaviors is moderately related to perceived team effectiveness and productivity. Person-focused behaviors were also found to be associated with perceived team effectiveness, productivity, and learning. Although the study results regarding team interdependence and their contribution to effectiveness, productivity, and learning lacked sufficient effect and sample sizes, the authors concluded that when task interdependence is higher, team leadership is “relatively” more important in achieving efficacious team performance outcomes. These findings support a SBT-based leadership curriculum in EM, because care of acutely ill patients has a high level of interdependent tasks.

During a resuscitation of a critically ill patient, the importance of leadership is paramount. The ability to maintain control of a resuscitation room through planning and preparation can be evaluated by looking for behaviors such as goal specification, evidenced by the leader’s act of summarizing the current understanding of a given situation. SBT encourages all participants to face decisions related to team processes and leadership challenges.

SBT to Enhance Leadership Training
Any leadership curriculum must have clear teaching goals with appropriate measurement tools to monitor the progress of its trainees. SBT assists in learning and assessment. The instructors must be clear in their vision of what constitutes an effective leader. The organizational psychology literature provides guidance for general best-practices in leadership, but real-world leadership experience in the ED is necessary to translate this into SBT scenarios. As debriefing is the key to any useful SBT experience, those leading the debriefing must be able to do more than merely review a scoring sheet of observed behaviors. They must also provide a clear connection between the concepts raised during the SBT, to the trainees’ actual practice. It is not enough to point out to a resident that they failed to “control the room” during a resuscitation. One must provide strategies, such as voice control, eye contact, positioning in the room, and task delegation, that can be employed to improve such control. This discussion is more challenging if the instructor has not had applicable leadership training or significant experience managing resuscitations. SBT is a means for initiating discussions and reflecting on performance, but debriefing activities related to leadership also need to be conducted after challenging clinical situations to reinforce key concepts.

Creating Leadership Scenarios
How does one create simulation scenarios to address leadership behaviors, permitting their measurement? As SBT can be absorbing and overwhelming, scenarios that are “clean,” i.e., do not provide multiple pathways for different outcomes, are more easily controlled to keep the teamwork and leadership issues at the forefront.

As noted, the SMARTER methodology provides a framework for scenario development that focuses on the desired KSAs in EM. Although not validated for leadership training purposes, it is reasonable to assume this system can be utilized for leadership-specific scenarios. As leadership is often manifested by decision-making, decisions the participants will face in the simulation should be defined. Scenarios that force
participants to address critical decisions or actions should be created. The scenarios should be crafted so that the patient’s progression provides additional triggers that elicit desired participant reactions. In most settings for leadership SBT, it is important not to make the outcome of the scenario dependent on the participants’ making the “right” decisions. It is possible that the team succeeds and the leader has fallen short of the desired training goals. To avoid unnecessary confusion in the learners, complex medical situations should only be used where the appropriate treatment is apparent. Immediately following the scenario, a teamwork debriefing should be conducted in which the focus is on the team leadership teaching points. The use of straightforward clinical cases minimizes digressions over the case’s appropriate management and boosts participants’ confidence in their own clinical abilities, permitting a less defensive, more open debriefing session. Measurement of the behaviors of interest requires that they be outlined prior to the case’s creation. An example observational checklist is provided in Figure 2. The more clearly defined the behaviors are, the more easily a measurement tool can be utilized. Tight control of the scenario’s progression increases the odds of generating the desired behaviors.

**Metrics for Leadership Assessment**

Once leadership behaviors have been defined, one can approach their metrics. This is the most challenging aspect of leadership training, teamwork training, and SBT in general. There are no universally applicable validated leadership measurement tools. Different settings require different skill-sets. Different types of teams have different goals, requiring distinctly unique metrics. What is measurable may not yield useful information. Examples of available measurement tools include the anesthesia nontechnical skills, which allows for individual member assessment and therefore may be used as a tool for leadership behavior assessment, compared to the communication and teamwork skills that evaluates whole team performance. More recently, an EM-specific system has been proposed as an event-based measurement approach. The SMARTER approach systematically links performance measures to ACGME core competencies. Each measurement tool applies to different arenas; none is universally applicable, especially in the evaluation of leadership skills. Measuring behavior often involves quantifying subjectivity; hence there is a need for metrics specific to a given situation and team member that can also permit standardized comparisons.

**CONCLUSIONS**

The health care community is interested in improving teamwork. Yet, numerous well-intentioned programs may fail to achieve the desired outcomes in team performance. We address the need for teamwork-focused SBT in EM and detail a process for its development. Teamwork is dynamic and complex, requiring an evaluation strategy that is dynamic as well. While no criterion standard metric or measure exists, there is a solid methodology for developing tools designed to assess the attributes sought. More research is needed to develop high-performing metrics. It is hoped that this discussion will assist those in EM, and the larger health care field, in the design and delivery of SBT for training and evaluating teamwork.

**References**


