Measuring Team Performance “In The Wild”: Challenges and Tips

Eduardo Salas
University of Central Florida

C. Shawn Burke
University of Central Florida

Jennifer E. Fowlkes
CHI Systems, Inc., Orlando, Florida

Since the 1980s, teams have risen to become one of the most dominant organizational strategies within industry and government (e.g., Mohrman, Cohen, & Mohrman, 1995; Osterman, 1994; Stewart, Manz, & Sims, 1999). This trend is only expected to continue as organizations increasingly operate in complex, dynamic environments that require coordinated adaptive action (Salas, Stagl, & Burke, 2004). Throughout history, work teams have evolved such that the types of teams used within organizations now run the gamut from temporary to permanent, small to large, real to virtual, self-managing to hierarchically led, colocated to distributed. In addition, organizational teams are increasingly interfacing with a variety of technologies and a cross-cultural workforce brought on by globalization.

The predominant reason for the tremendous rise in the use of teams within organizations has been the mythical assumption that teams will automatically result in a competitive advantage for the organization by producing better outcomes more efficiently. As an illustration of this mind-set, Proctor and Gamble (a pioneering company in the use of teams) felt so strongly about their implementation of teams they considered their use a trade secret (Stewart et al., 1999). Despite the enthusiasm over teams, history and real-world examples have shown that teams do not always work well (Hackman, 1998). Moreover, there are surprisingly few mechanisms in place within organizations to system-
attractively diagnose team effectiveness. As a result, managers must often rely on informal data to determine what goes wrong with teamwork, an unfortunate consequence given the complex nature of teamwork. Thus, to support the diagnostic assessment of teamwork, team performance measurement systems are needed (Kendall & Salas, 2004).

Though we have learned much over the past 20 years concerning the characteristics of effective team performance measurement systems (see Table 10.1), most of this work has been conducted in the laboratory and has not yet filtered into the measurement of teams “in the wild”—teams performing in complex and dynamic settings. Although there are potentially several reasons for this lack of transition, we briefly discuss two. First, as researchers and academics we do not do a good job of marketing and translating what we have learned about teams to practitioners. Second, even when we do market or translate, we often fail to realize the additional hurdles that must be overcome within organizations working with real teams. Therefore, the purpose of this chapter is fourfold. First, we briefly describe a subset of teams that are prevalent in organizations today, yet often neglected within laboratory settings. We do this to provide a concrete example of the types of teams being seen in organizations. Second, we highlight some of the challenges facing those tasked with measuring the performance of such teams. Next, we briefly describe how some of these challenges are currently

<table>
<thead>
<tr>
<th>Measurement Characteristics</th>
<th>Why It's Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Measurement needs to assess relevant team competencies (Cannon-Bowers &amp; Salas, 1997).</td>
<td>• Teamwork is a complex entity comprising behavioral, cognitive, and affective components. Team performance models can guide the practitioner in the competencies to assess, but models are plagued with labeling inconsistency and confusing terminology. They can't be used in isolation.</td>
</tr>
<tr>
<td>• Measurement needs to assess performance of teams, subteams, and individuals (Klein &amp; Kozlowski, 2000).</td>
<td>• The underlying nature of each particular team will dictate the relative importance of particular teamwork dimensions.</td>
</tr>
<tr>
<td>• Although members are interdependent, teams are composed of individual members (Dickinson &amp; McIntyre, 1997).</td>
<td>• Nesting is common within teams. Individual members are members of an interdependent team that is often a member of a higher team and may also contain subteams. Each level will have an impact on adjacent levels.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Measurement Characteristics</th>
<th>Why It’s Important</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Measurement needs to assess process as well as outcome (Cannon-Bowers &amp; Salas, 1997).</td>
<td>• Team competencies exist at several levels (e.g., individual competencies, team competencies held at the individual level, team competencies held at the team level; Cannon-Bowers &amp; Salas, 1997). The level at which measurement should collect information should be driven by theory.</td>
</tr>
<tr>
<td>• Measurement needs to control task content (Dwyer &amp; Salas, 2000).</td>
<td>• Outcome measures provide an indication of what happened but not why it happened. Though fairly easy to obtain, outcome measures often contain variance not attributable to the team itself.</td>
</tr>
<tr>
<td>• Measurement should facilitate observation (Baker &amp; Salas, 1997).</td>
<td>• Process measures are those that provide a means to capture the moment-by-moment behaviors exhibited by teams, thereby allowing diagnosis of team strengths/weaknesses (Dwyer &amp; Salas, 2000).</td>
</tr>
<tr>
<td>• Measurement should utilize multiple method approaches.</td>
<td>• Organizations that do not collect both process and outcome measures often end up with incomplete and potentially misleading pictures of team performance. For example, a team can produce a correct decision (outcome) yet the social interactions that occur in reaching that decision (the process) may be detrimental to future working relationships.</td>
</tr>
<tr>
<td></td>
<td>• A key characteristic of any team performance measurement system is the extent to which what is being practiced or trained can be determined. Controlling task content guarantees that the competencies one wishes to target will be given the chance to be exhibited and thereby captured.</td>
</tr>
<tr>
<td></td>
<td>• Important due to the multidimensional nature of teamwork.</td>
</tr>
<tr>
<td></td>
<td>• Provides an objective standard that can be compared across teams and members.</td>
</tr>
<tr>
<td></td>
<td>• Process measures rely almost exclusively on observation.</td>
</tr>
<tr>
<td></td>
<td>• Teams of observers are often needed.</td>
</tr>
<tr>
<td></td>
<td>• Observers may have different areas of expertise if subteams are contained within the team.</td>
</tr>
<tr>
<td></td>
<td>• Measurement techniques may be differentially suited to capture the behavioral, cognitive, or affective aspects of teamwork.</td>
</tr>
<tr>
<td></td>
<td>• Helps to prevent same source bias.</td>
</tr>
<tr>
<td></td>
<td>• Measurement can be directed at reactions, learning, skills, or organizational impact (captured by different methods). Each area provides different information culminating in a picture of effectiveness.</td>
</tr>
</tbody>
</table>
being handled. Finally, we conclude by providing some guidance as to how practitioners can overcome some of the challenges associated with team performance measurement.

WORK TEAMS IN THE WILD

Prior to enumerating the challenges, we must first understand the entity that we are talking about. Teams have been defined as a set of two or more individuals interacting adaptively, interdependently, and dynamically toward a common and valued goal (Salas, Dickinson, Converse, & Tannenbaum, 1992). Whereas the types of teams seen within laboratory environments can be fairly narrow, those that appear in the wild are broad and diverse. As such, we next provide a brief description of a subset of work teams that we feel offer great challenges for the measurement of team performance. The subset that we chose to focus on were those teams that we felt were most highly impacted by global competition and technological advances (i.e., in many cases those most different from work teams of old).

Distributed Teams

Organizational downsizing has left a variety of expertise to cover global operations in which mergers have created a patchwork of mismatched skills (Joy-Matthews & Gladstone, 2000). Technological advances have provided one solution to this problem—distributed teams. Though it was once the case that the interdependent nature of teams required all teams to be colocated, technology has banished this requirement. Distributed teams have been defined as, “groups of geographically and/or organizationally dispersed coworkers that are assembled using a combination of telecommunications and information technologies to accomplish an organizational task” (Townsend, DeMarie, & Hendrickson, 1996, p. 17). These distributed teams are often large in size and composed of members with different areas of background and expertise. Bell and Kozlowski (2002) have further differentiated among types of distributed teams that are being seen within organizational settings. Specifically, their typology argues that distributed teams vary along a set of four characteristics: (a) temporal distribution (entrained by real time or distributed across time), (b) boundary spanning (permeability of these boundaries once crossed depends on task complexity; functional, organizational, cultural), (c) life cycle (discrete, continuous), and (d) member roles (members often hold multiple roles within and across virtual teams).
Human–Computer Teams

Technological advancements have also driven the evolution of human–computer teams. With the increase of virtual reality, simulation, and electronic support tools, it is increasingly common for fellow team members to be synthetic (e.g., virtual reality, simulation) or machines (e.g., electronic support tools, robots, automation). The term synthetic team members refers to those team members that are computer generated. Technology has made it possible to link live forces with virtual simulation forces through the use of instrumentation. The use of synthetic team members has applications not only for the training of colocated teams, but is often used in distributed mission training (e.g., Bennett & Crane, 2002; O’Neil & Andrews, 2000). Although the use of synthetic team members is most commonly seen within military training contexts, their use is beginning to filter into mainstream organizations.

Perhaps currently more common in an operational environment is the use of machines as team members (i.e., electronic performance support systems, robots, intelligent tutoring systems). Electronic performance support systems (EPSSs) have been described as an integrated electronic environment that facilitates just in time training by being easily accessible and structured to provide immediate, individualized access to information, guidance, advice and assistance, data images, tools, and assessment monitoring systems to facilitate job performance with minimal intervention by others (Bastiaens, 2002; Gery, 1991). In addition, EPSSs often reduce the complexity or number of steps required and provide a decision support system that assists the human in the identification of the appropriate action for a particular set of circumstances. Elliott (2000), for example, describes an intelligent agent technology that functions as an EPSS by providing real-time decision support to weapons directors within an airborne warning and control system (AWACS) environment.

EPSSs can be conceptualized as forming one end of the human–computer team in a number of domains. For example, within the instructional domain, an intelligent tutoring system (ITS) can be viewed as an EPSS as it mimics the learning and diagnosis that would normally occur between a student and a teacher. Specifically, the ITS serves as a team member by diagnosing the student’s performance based on a comparison between an a priori defined expert model and the learner model. This information is then forwarded to a pedagogical model that makes decisions about the type of feedback and tutoring to be offered. Another large user of EPSS is the aviation community’s use of adaptive automation. Adaptive automation refers to the idea that automation can be turned on/off by either the computer or the human depending on situational needs. There are also a wide variety of companies using EPSSs based on the premise that, instead of spending tremendous amount of upfront time to train people on skills they might not use right away, they can train people just-in-time with the use of EPSSs. Although the EPSS can function as an electronic librarian,
expert adviser, patient tutor, and administrative assistant, it's most likely viewed as a
team member when serving in the adviser or tutor role due to the interdependence
present.

Finally, one last instance of human-computer team members is the increasing
use of robotics and automation to help complete the more automatic and mundane
tasking in several organizational jobs. Many organizations are increasingly relying
on robotics for work such as that conducted along assembly lines. Additionally, ro-
botics can also be used to help extend the reach of human team members in vari-
ous environments (e.g., search and rescue, reconnaissance functions). In terms of
automation, the aviation community has been one of the largest users and has iden-
tified several lessons learned concerning human interaction with automated sys-
tems (Olson & Sarter, 2000; Sarter & Amalberti, 2000).

High-Performance Teams: Aviation/Medical

The complexity that has been brought forth by technology and globalization
has caused an increase in high-performance teams (HPTs). These teams ac-
complish mission-critical tasks in the face of complex and dynamic environ-
ments where mistakes can easily result in highly publicized loss of life. More-
over, HPTs are most indicative of ad hoc or just-in-time teams and members
often have distributed expertise. In the case of aviation crews, the composi-
tion of both the flight crew and cockpit crew is often varied not only across
flights, but sometimes within different legs of the same flight. A similar situ-
ation exists within the medical community's operating room and trauma teams.
The situation is further complicated within these latter teams in that the equip-
ment that they are working with may be different from hospital to hospital, as
well as across operating rooms within the same hospital. Therefore, not only
are they working with a team whose composition may often change and tends
to be hierarchically structured with clear lines of authority, but the equipment
may also change. Standard operating procedures become very important
within HPTs due to the complexity present and ad hoc nature.

Self-Managed Teams

Self-managed teams (SMTs) have also become prevalent in the workforce. De-
spite their occurrence in operational environments, SMTs are not often repre-
sented within a laboratory context. Self-managing teams have been defined as
teams with a relatively flat structure whose members are expected to manage
themselves and focus on improving their own work processes (Stewart et al.,
1999). SMTs are typically afforded high levels of empowerment and, although
responsible for much of what would traditionally be accomplished by leader-
ship outside the team, they are subject to external control in the sense that they
do not set strategic objectives. The following characteristics have been identi-
fied as being indicative of SMTs: (a) distinct, recognizable tasks that workers can identify with, (b) high levels of empowerment (e.g., discretion over how work is done, scheduling, and assigning tasks), (c) members with a variety of task-related skills, (d) flat structure with shared leadership in the form of facilitation or coaching, (e) members who rotate between various scheduling and coordinating roles within the team, and (f) compensation and performance feedback for the group as a whole (e.g., Hackman, 1987; Stewart et al., 1999; Wellins, Byham, & Wilson, 1991). In addition, members of SMTs may experience role ambiguity due to the fluid nature of roles within these teams as well as additional stress from the fluidity present within these teams.

Summary

As technology and global competition continue to increase, organizational teams evolve to meet the internal and external demands driven by these changes in order to remain competitive within their niche. Whereas in some cases this evolution consists of minor adaptations, in other cases, it results in the formation of new team types (i.e., human–computer, distributed teams). Therefore, the teams that are presently operating in the wild are in many ways not only different from the ad hoc teams often composed in laboratory settings, but also from organizational teams of the past. For example, the following have been mentioned as characteristics of teams currently operating in the wild: (a) ill-structured, ambiguous tasks, (b) distributed (physically and/or functionally), (c) heavy workload, time pressure, and high consequences for error, (d) communication modes that are increasingly electronic and intensive, (e) a mixture of human–human teams and human–computer teams, (f) multicultural (organizational and national level), and (g) multiteam systems composed of ad hoc teams. In addition, teams that operate in dynamic environments often have to contend with information overload and stressful environments. The aforementioned characteristics pose several challenges not only in regard to the competencies (knowledge, behavior, and attitudes) that need to be trained and measured, but also in terms of the actual implementation of such measures (see Table 10.2). Next, we discuss several of these challenges.

MEASUREMENT CHALLENGES IN THE WILD

Whether conducted within the laboratory or operational setting, team performance measurement refers to the systematic assessment of teamwork and taskwork competencies as well as the assessment of team outcomes. This assessment can serve a variety of purposes within organizations, ranging from research to practical purposes such as certification/assessment, problem diagnosis, selection, remediation, and evaluation of training. However, due to the complexity present
<table>
<thead>
<tr>
<th>Characteristic/Work Team</th>
<th>Effect on Team</th>
<th>Task Performance Consequence</th>
<th>Team Performance Measurement Requirements</th>
</tr>
</thead>
</table>
| Large Number of Team Members | • Reduced capability for flexibility  
• Heterogeneous subteams | • Less ability for online planning/replanning  
• Management of heterogeneous subteams | • Assessment of preplanning, including contingency planning  
• Use of implicit coordination during task execution (i.e., shared mental models)  
• Information management (how is information distributed before, during, and after task performance)  
• Understanding of/knowledge of subteams  
• Assessment of intra- and interteam cooperation/processes  
• Task communication that builds situational awareness  
• Implicit coordination during task execution (i.e., shared mental models) | |
| Distributed Team Members | • Loss of visual cues  
• Restricted information flow | • Degraded communications  
• Fewer communications | • Task communication that builds situational awareness  
• Use of standardized communication formats  
• Use of implicit coordination during task execution (i.e., shared mental models) | |
| Electronic Team Members | • Lack of immersion  
• Greater amounts of information  
• Information may be filtered  
• Automation | • Reduced situational awareness  
• Harder to develop cohesion  
• Social interaction may differ  
• Fewer social cues/harder to interpret actions  
• Degraded team and team interaction shared mental models  
• Harder to develop cohesion  
• Harder to develop trust  
• Less communication with entities outside the team  
• Susceptibility to groupthink | • Built-in task cues to prompt other team members  
• Teamwork processes such as backup and monitoring  
• Trust  
• Complacency with automation or electronic member  
• Misuse or disuse of automation or electronic member  
• Transportable team skills  
• Standard operating procedures  
• Clear roles  
• Situation awareness  
• Process feedback  
• Real-time measurement  
• Self-leadership skills/self-correction skills  
• Mutual monitoring |
within the operational setting and the teams embedded within this setting, practitioners charged with developing and implementing such systems will face a variety of challenges. Next we briefly describe what we feel are the top six challenges based on our experience.

**Challenge 1: What to Measure**

There is no getting around the fact that teams are complex dynamic entities. Although there are a variety of models and taxonomies developed by the research community that practitioners can use to guide them in the determination of the appropriate competencies required for effective teams (e.g., Argote & McGrath, 1993; Campion, Papper, & Medsker, 1996; Fleishman & Bocco, 1992; Tannenbaum, Beard, & Salas, 1992; West, 1996), when used in isolation these models pose several problems. These models are often not practical to use for practitioners because labels given to teamwork dimensions are plagued with inconsistency and confusion (e.g., many models use the same labels to refer to different constructs) (see Burke, Cannon-Bowers, Volpe, & Salas, 1993). Furthermore, all teams are not created equal. Teams come in a variety of shapes, sizes, and configurations, have differing levels of interdependence, and are faced with a variety of contextual factors that impact team process and the resulting performance (Sundstrom, McIntyre, Halfhill, & Richards, 2000). The underlying nature and requirements of each particular team will dictate the relative importance of the various teamwork dimensions (McIntyre & Salas, 1995). Therefore, in order to ensure that team performance measurement is assessing competencies relevant to a particular team, the nature of each team and its operating circumstances need to be considered. However, the predominant number of models do not explicitly acknowledge the fact. This combined with the inconsistent terminology across models often leaves the practitioner confused when attempting to use theory to guide the development of team performance measurement systems.

**What Is Being Done?**

There are at least two streams of research that can be used to aid organizations to determine the correct competencies to assess: (a) team-type taxonomies (e.g., Devine, et al., Sundstrom et al., 2000) and (b) team-competency taxonomies (e.g., Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995). There has been much work conducted on the types of teams that appear within organizations. Most recently work by Sundstrom et al. and Devine identified several team types as delineated by various contextual variables. Sundstrom and colleagues developed a taxonomy consisting of kinds of work teams, distinguished by the type of work they accomplish: (a) production, (b) service, (c) management, (d) project, (e) action/performing, and (f) parallel teams. Furthermore, each of these teams differs in at least four key factors: (a) level of
authority within the organization, (b) time until the team is disbanded, (c) degree of specialization, independence, and autonomy in relation to other work units, and (d) degree to which members are interdependent within the team as well as with forces outside the team.

Devine (2002) developed a team-type taxonomy consisting of 14 team types that can be categorized according to seven contextual variables (i.e., task structure, active resistance, hardware dependence, health risk, fundamental work cycle, physical ability, and temporal duration). Devine’s taxonomy sought to refine existing taxonomies by eliminating the confounding of attribute dimensions and superordinate team dimensions, integrating new team types, and expanding on critical contextual variables that highlight key differences in team effectiveness across team types.

Team-type taxonomies, like those just described, can help the practitioner determine the type of team that she or he is working with. Once this is done, matching the relevant contextual variables with the work that Cannon-Bowers et al. (1995) has done can help determine the appropriate competencies to be trained in a practical, theoretically driven manner. Specifically, Cannon-Bowers and colleagues argued that the competencies related to effective team performance may vary along two axes (i.e., team/task, generic/specific). Team-generic competencies are those held by individual members that can influence team performance regardless of the particular team members involved (e.g., communication skills). As such, these skills are transportable and applicable across teams. Conversely, team-specific competencies are held by individual team members, but influence performance only with regard to the team that the individual is currently working within. On the flip side, competencies will also vary along the same two axes with regard to the task itself. Task-generic competencies are those that will influence team performance across all tasks (e.g., planning skill), whereas task-specific competencies are tied to promoting effective performance on a particular task or context. Cannon-Bowers et al. then used this information to create four broad categories of competencies: context-driven (team/task-specific), team-contingent (team-specific/task-generic), task-contingent (team-generic/task-specific), and transportable (team/task-generic). Finally, the authors delineated the knowledge, skills, and attitudes needed by each of the four categories (see Table 10.3).

Working from the team-type taxonomies and using the contextual variables identified, one should be able to begin to match each team type with one of the four categories contained within Cannon-Bowers and colleagues’ (1995) work containing a delineation of knowledge, skills, and attitude (dependent on which of the four quadrants that team is closest to). Cannon-Bowers and colleagues also offer the following guidance to practitioners. First, context-driven competencies are generally required when teams are faced with a situation in which the task is highly demanding and requires members to quickly adapt strategies. Membership in teams requiring context-driven competencies is relatively stable.
<table>
<thead>
<tr>
<th>Nature of Team Competency</th>
<th>Description of Team Competency</th>
<th>Knowledge</th>
<th>Skills</th>
<th>Attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context-driven</td>
<td>Team-specific, Task-specific</td>
<td>Cue-strategy associations</td>
<td>Task organization</td>
<td>Team orientation (morale)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Task-specific teammate characteristics</td>
<td>Mutual performance monitoring</td>
<td>Collective efficacy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Team-specific role responsibilities</td>
<td>Shared problem-model development</td>
<td>Shared vision</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Shared task models</td>
<td>Flexibility</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Team mission, objectives, norms, resources</td>
<td>Compensatory behavior</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Information exchange</td>
<td></td>
</tr>
<tr>
<td>Team-contingent</td>
<td>Team-specific, Task-generic</td>
<td>Teammate characteristics</td>
<td>Conflict resolution</td>
<td>Team cohesion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Team mission, objectives, norms, resources</td>
<td>Motivation of others</td>
<td>Interpersonal relations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Relationship to larger organization</td>
<td>Information exchange</td>
<td>Mutual trust</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intrateam feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Compensatory behavior</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Assertiveness</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Planning</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Flexibility</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Morale building</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cooperation</td>
<td></td>
</tr>
<tr>
<td>Nature of Team Competency</td>
<td>Description of Team Competency</td>
<td>Knowledge</td>
<td>Skills</td>
<td>Attitudes</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Team role–interaction patterns</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Procedures for task accomplishment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accurate task models, Accurate problem models</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boundary-spanning role, Cue–strategy associations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportable</td>
<td>Team-generic, Task-generic</td>
<td>Teamwork skills</td>
<td>Morale building, Conflict resolution, Information exchange, Task motivation, Cooperation, Consulting with others, Assertiveness</td>
<td>Collective orientation, Belief in importance of teamwork</td>
</tr>
</tbody>
</table>

and the range of tasks performed by the team is relatively small (e.g., combat teams, sports teams). Second, team-contingent competencies are generally required when team membership is stable, but members must work together across a wide range of tasks (e.g., self-managing teams, quality circles). Third, task-contingent competencies are generally required when members perform a small range of tasks, but team membership is not constant (e.g., air crews, medical teams). Finally, transportable competencies are generally required by teams who work on a variety of tasks with a variety of team members (e.g., task forces, project teams).

Challenge 2: How to Ensure Reliability, Diagnosticity, and Practicality All at Once?

This is perhaps the biggest challenge resulting from moving performance measurement work from the laboratory to the wild. The laboratory has taught those developing and implementing measurement instruments that they must be reliable (i.e., capture a stable characteristic of the team), sensitive (i.e., reflect measurable changes in the characteristic of interest), and valid (i.e., capture meaningful aspects of performance). In addition, researchers have identified several other requirements and guidelines for the development and implementation of team performance measurement systems (see Table 10.1). However, when practitioners take these requirements and guidelines and apply them within actual organizations they face many challenges. For example, within the laboratory environment, there is often the luxury of time (in terms of both development and implementation). However, within actual organizations, resources are usually tight (e.g., money, time). In addition, for those charged with team measurement in the wild, there is the requirement that measures have face validity (i.e., be accepted by those completing the measures) and be easy to implement. Often measures of team process and performance are cumbersome in both development and implementation. Team performance measurement systems implemented within organizations should be as unobtrusive as possible, yet much of our common stock and trade of team performance measurement instruments are based on self-report or observation.

What Is Being Done?

Though striking a balance between the development and implementation of measurement systems that are theory driven, easy to administer and make sense of, and possess face validity is a definite challenge, there are a number of efforts under way to accomplish just that. A great part of this process is including subject matter experts from the operational environment in the measurement development process as well as those representing measurement development. Additionally, work is currently being conducted to make team performance measurement instruments
less intrusive, easier to implement, and better able to provide real-time feedback to users (e.g., Fowlkes, Lane, Salas, Franz, & Oser, 1994; Kiekel, Cooke, Foltz, Gorman, & Martin, 2002). A number of these developments are further elaborated upon within the specific challenges that follow.

Challenge 3: How to Capture Dynamic Performance

Another challenge that presents itself with regard to the assessment and diagnosis of teams within naturalistic settings is the need for timely feedback. Unlike laboratory teams, teams in the wild often do not have the luxury of time—every second counts. Organizational teams operate under several time constraints that serve to make the importance of real- or near real-time assessment and diagnosis imperative to remaining competitive in a dynamic, complex global market. Each moment they are not performing to their potential costs the organization in terms of lost productivity and resources. Furthermore, the longer members persist in incorrect or maladaptive teamwork behaviors the harder it is to correct their knowledge base (i.e., mental models) with regard to the proper actions.

What Is Being Done?

Although this remains a challenge for measurement systems in the wild, there has been progress made within the last few years. The challenge here lies not so much within the dynamic assessment of outcome type data (as technology has helped in this arena by allowing the dynamic capturing of task related actions that can be incorporated into feedback mechanisms given to the team), but in the capturing of process data. There are primarily two techniques that are being used to assist in providing members or team leaders with near real-time assessment information: (a) observation and (b) automated performance measurement. Though acknowledging the fact that observational methods of assessing team performance do not represent real-time assessment, if structured carefully, they may be used to facilitate near real-time assessment and diagnosis. One tool that has been used successfully to provide near real-time feedback is the event-based assessment technique (EBAT). Using this technique, behavioral checklists are created that identify actions that should be associated with key a priori determined events. Observers use these checklists to aid them in the identification of behaviors of interest, as well as the delivery of concise and timely feedback (see Fowlkes et al., 1994; Fowlkes & Burke, 2005; Smith-Jentsch, Zeisig, Acton, & McPherson, 1998). PDAs (personal digital assistants) have also been used to house these checklists and help facilitate the transfer and combining of information (see also Elliott, Dalrymple, Regan, & Shiflett, 2001, for a similar approach).
On the automated side of the house, researchers at the Air Force Research Laboratory working within a distributed mission training environment have developed a competency-based, embedded performance measurement system for tracking individual and team performance in an effort to provide near real-time feedback to pilots (Schreiber, Watz, & Bennett, 2003). This proof-of-concept system resides on the network, listens to network traffic, and collects appropriate variables at a specified rate as defined by mission-essential competencies. Mission-essential competencies are “higher order individual, team, and inter-team competencies that a fully prepared pilot, crew, or flight requires for successful mission completion under adverse conditions and in a non-permissive environment” (Colegrove & Alliger, 2002, p. 12-2). It then uses measurement algorithms to input and output variables (i.e., data) in several data formats for either feedback or research purposes (see also Bennett & Crane, 2002; Castillo et al., 2002).

Finally, intelligent tutoring systems may also support real-time assessment when they are configured to provide real-time coaching. For example, the Navy has embedded an ITS within a virtual environment (COVE—conning officer virtual environment) in an effort to reduce the workload on the instructor (Cohn & Patrey, 2001). This ITS dynamically coaches the student through several difficult maneuvers required by the conning officer (ship’s driver).

Challenge 4: How to Capture Communication

Communication has been defined as, “the process by which information is clearly and accurately exchanged between two or more members” (Salas & Cannon-Bowers, 2000, p. 317). The assessment of team communication is probably one of the most researched aspects of team process as it often serves as the mode through which other process measures are gleaned. Within natural environments the collection of team communication data, though vital to the assessment of a team’s effectiveness, becomes a major challenge in two ways. The first challenge in capturing this information applies across all organizational teams and relates back to the issue of real-time assessment. Specifically, though communication data can be captured relatively unobtrusively via automated recording, extracting meaningful data from this information is a notoriously laborious process that may involve transcription of the communication as well as actually coding the information to identify key behaviors or communication sequences. The challenge here lies within the creation of smart and adaptive systems that will respond to the dynamic nature of team performance, as well as the creation of tools by which to make sense of the data gathered so that it can be fed back to teams in real time.

The second challenge relates to the use of observational methods by which to measure team communication (e.g., event-based checklists). This method relying almost exclusively on observation poses challenges for several types of teams op-
erating in the wild. For example, distributed and/or large teams may require more resources in order to capture communication as members are not face-to-face and one observer may not have access to the communication of all. This in turn requires a team of observers whose data must be weighted and combined. Additionally, Weisband and Atwater (1999) argue that ratings of team performance may vary depending on whether ratings take place within a face-to-face environment or distributed environment.

Teams that use digital or electronic modes for communication also pose challenges. It is harder to “make sense” out of communication when the form of communication becomes digital or electronic as compared to the participant’s actual voice, because verbal cues accompanying syntax often help observers in their determination. The mode of communication may add other challenges such as those regarding sequence or timing of communication actions and requests when there are delays in communication due to technology.

What Is Being Done?

Within actual organizations, the assessment of team communication remains painful and laborious. However, increases in technology related to the automatic capturing of communication, and in some instances overlying this on behavioral sequences, may aid with some of the difficulties. Perhaps the most promising recent development has been the use of latent semantic analysis applied to team communication data. Latent semantic analysis is “a machine learning method for automatically extracting and representing knowledge in massive databases of electronic text” (Castillo et al., 2002, p. 4; see also Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990). As a database must be built from which the machine can “learn,” this method still requires arduous work on the front end, but considerably reduces time on the back end. This method has been extensively validated in both laboratory and field tests, and when applied to team communication, has been argued to show promising results as a surrogate measure of team cognition (Kiekel et al., 2002; Landauer, Foltz, & Laham, 1998).

Challenge 5: How to Capture Team Cognition and Be Practical

Organizational teams are increasingly interacting in dynamic complex environments where technology has afforded members the ability to receive an amount of information that can overload members’ cognitive resources if not managed properly. In addition, work teams are increasingly becoming culturally diverse and members may operate with different values and assumptions that serve to guide behavior. All of these factors combine to make the measurement of team cognition a vital component in determining the underlying reasons for why a team succeeds or fails. Although the measurement of cognitive processes and
states (i.e., shared mental models, metacognition, situational awareness) has been recently touted by many as being important within teams (i.e., Cannon-Bowers & Salas, 1998; Milton, Shadbolt, Cottam, & Hammersley, 1999), its measurement poses several challenges for teams operating in the wild. First, the construct that one is attempting to measure is not observable, it is in team members’ minds. As such, one challenge is how to elicit the information needed to measure this construct without imposing one’s own mental model on participants. Second, within an organizational setting, how can this be accomplished in a practical manner, as most measurements of team cognition are time consuming and obtrusive?

What Is Being Done?

The measurement of team cognition may be one of the more difficult challenges for those in an operational environment as researchers are still working on developing many of the procedures for measurement within laboratory settings (and even those provide many challenges). Although there is still much progress to be made, we next present a brief look at how two key constructs that represent aspects of team cognition are currently measured.

Shared Mental Models. Shared mental models refers to the degree to which team members share similar mental models (cognitive representations) with regard to the team and its objectives, team roles, behavior, and interaction patterns (Cannon-Bowers, Salas, & Converse, 1993). Shared mental models have been argued to be composed of knowledge, behaviors, and attitudes (Kraiger & Wenzel, 1997). As such, measurement techniques have typically focused on measuring a subset of the following: how information is processed, structural knowledge, common attitudes, or shared expectations. In terms of the methods most commonly used within organizations, most measurement has been done within the context of training. Techniques most commonly used include Pathfinder, card sorts, and self-report questionnaires. Pathfinder is a quantitative multidimensional scaling technique that produces a pictorial representation of the linkages between targeted concepts, as well as providing quantitative measures of the internal consistency of these representations and the similarity between the representations of dyads (Schvaneveldt, 1990). Team members provide the input for this technique by rating the relatedness of constructs designated by the person responsible for measurement. Pathfinder can be delivered via a computer or paper-and-pencil. Related to Pathfinder is PRONET (Cooke, Neville, & Rowe, 1996). PRONET is a sequential analysis tool that relies on Pathfinder to graphically determine what events typically follow one another for a given time lag. Card sorts are another way of capturing participant knowledge by requiring members to sort piles of predefined concepts into piles based on how they believe the concepts are related. Recently the Team Performance Labo-
ratory at the University of Central Florida has developed software (e.g., TPL-KATS, © 2001) that facilitates the administration of the card sort task by automating the process. Finally, self-report questionnaires are used. The structure of these questionnaires varies dependent on the type of knowledge that is captured. Mohammed, Klimoski, and Rentsch (2000) provide an evaluation of much of this work.

**Situational Awareness.** Situational awareness (SA) has been defined as the “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1988, p. 97). The assessment of this construct becomes most important within distributed, self-managed teams, or those operating within fast-paced environments. These teams are likely to have the most difficult time maintaining awareness due to either being dispersed, being isolated, and/or the dynamism involved in the setting.

Though team SA is just beginning to be measured in the wild, measures for individual SA have been repeatedly used in the wild. Cooke, Stout, and Salas (2001) have begun to provide a methodology for measuring team SA by the modification of existing methods of knowledge elicitation (see Table 10.4). When measuring team SA, Cooke and colleagues argue that there are four primary targets: (a) situation models, (b) mental models, (c) cues in the environment and the situation assessment process, and (d) team process behaviors. It is important to note at this point that shared mental models have been argued to be a precursor to team SA (Cooke et al., 2001). Perhaps the most predominant measurement tool that has been used in the wild is the Situational Awareness Global Assessment Technique (SAGAT). SAGAT uses simulation freezes and SA queries to obtain direct and objective measures of SA. The basis of this technique lies within the fact that operators are temporarily stopped at various points during their task (real or simulated) and are momentarily queried on the current SA. Queries are normally specific, open-ended, and devised to tap the completeness and accuracy of the situation model, not its contents. This technique still needs validation to show that it predicts performance.

**Challenge 6: How to Assess Performance in a Culturally Diverse Team?**

Technological advances (i.e., the enabling of distributed teams) that allow increased global competition will make it more likely that teams will be composed of members from different cultures (organizational, national). This trend is being seen in a wide range of companies (e.g., U.S. companies operating abroad, U.S. companies operating within the United States, and in virtual or just-in-time organizations where multiple companies band together for a short time). This trend must be taken into account, for cultural diversity will likely impact not only the
team's cohesion, communication, and coordination, but the cognitive frame that members bring to bear on problems and interactions. In turn, this cognitive frame affects how members respond to subjective measurement tools. Gelfand, Raver, and Holcombe (2002) cite numerous methodological challenges that exist when conducting cross-cultural organizational research, many of which can be applied to the development and implementation of measurement systems. Things that must be considered include, but are not limited to the following: (a) the appropriateness of the method for all cultures being evaluated, (b) the reliability of the measurement system, and (c) the ethical acceptability of the method. In addition, there may be challenges in terms of the interpretation of content, the willingness to respond to certain types of questions or those posed in certain formats, and the willingness to accept the results of evaluations (i.e., feedback), depending on the manner in which it was conducted and by whom. As an example of some of the aforementioned challenges, Moshinsky (2000; as cited in Gelfand et al., 2002) reported that despite the instructions to work independently, Russian participants worked collaboratively to answer a survey. As a group, they were reading the questions aloud, deciding upon a group answer, and

### Table 10.4
The Measurement of Team Situation Awareness

<table>
<thead>
<tr>
<th>Target</th>
<th>Assessment Methods</th>
<th>Assessment Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Situation Models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Individual</td>
<td>Queries</td>
<td>Query accuracy</td>
</tr>
<tr>
<td>• Team</td>
<td>Team orientated queries and intrateam comparison of query responses</td>
<td>Intrateam query accuracy and intrateam similarity</td>
</tr>
<tr>
<td>2. Mental Models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Individual</td>
<td>Compare knowledge structure to individual referent</td>
<td>Accuracy based on similarity to referent</td>
</tr>
<tr>
<td>• Shared</td>
<td>Compare team knowledge to team referent and intrateam comparison of knowledge structures</td>
<td>Team accuracy based on similarity to team referent; intrateam similarity</td>
</tr>
<tr>
<td>3. Cues and Situation Assessment (Individual)</td>
<td>Tracking eye movement</td>
<td>Presence or absence of eye movements associated with a cue</td>
</tr>
<tr>
<td></td>
<td>Policy capturing</td>
<td>Regression weights associated with cues</td>
</tr>
<tr>
<td></td>
<td>Experimental manipulation</td>
<td>Differences in task performance with manipulation</td>
</tr>
<tr>
<td>4. Team Process Behaviors (Team)</td>
<td>Compare team process representation to team referent</td>
<td>Behavioral protocol, effectiveness based on similarity to team referent</td>
</tr>
</tbody>
</table>

*Note. From Cooke et al. (2001). Copyright 2001 by Human Factors and Ergonomics Society. Adapted by permission.*
all circling the same response. These participants found the individual survey methodology to be inconsistent with their cultural experiences and values, and therefore, modified the instructions to the more culturally appropriate group consensus task.

Those charged with developing such measurement and feedback systems need to be aware of these differences and how they might impact their tools (see Salas, Burke, Fowlkes, & Wilson, 2004). Practitioners need to be careful with regard to not only the method chosen for performance assessment, but also the wording contained within the actual tool because it might be interpreted differently depending on the individual culture involved. In terms of the feedback delivered to aid remediation, culture may also play an important role in that each culture has its own norms that may interact with the way in which feedback is delivered, as well as who should deliver feedback. The importance of this is that in order for performance assessment to have a long-term impact on the organization, the diagnosis and suggested remediation efforts must be accepted by the respective teams. The bottom line is that practitioners should not assume that measurement systems developed for use in a specific national culture will be interpreted the same in different cultures. The aviation community has already found this out in terms of their crew resource management training—cultures interpret the training differently (Helmreich & Merritt, 1998). Additionally, the degree to which emotion is expressed in feedback must be aligned with cultural perceptions or else it may be misinterpreted.

What Is Being Done?

The measurement of multicultural teams or groups is a fairly new endeavor. Though diversity has gained increased attention in our society, there has not been much done at the level of team measurement where there is cultural heterogeneity within a team. However, in terms of the challenges and preliminary guidance in overcoming these challenges, one can begin to look at the literature on cross-cultural diversity because much has been done at an individual level. Relevant aspects of this work need to be extracted and extrapolated in terms of how it translates to team performance measurement.

SOME PRACTICAL MEASUREMENT TIPS FOR ASSESSMENT IN THE WILD

Tip 1: Make Sure Your Team Is Really a Team. As noted, teams come in all sizes and shapes—they may be temporary/permanent, members may be human/computer, members may be colocated/distributed, and so forth. Therefore, perhaps the starting point in the design and application of a quality team performance measurement system is ensuring that your team is really a
team (see Salas et al., 1992, for a definition). In addition, members are each assigned specific roles or functions to perform and the team has a limited life span. As teams continue to evolve into many different forms, the decision as to whether the entity is truly a team may become confusing. However, as long as practitioners keep the basic defining qualities in mind—interdependence (see Saavedra, Earley, & Van Dyne, 1993, for description of degrees of interdependence), common goal, limited life span, and specific roles—these qualities should serve to guide them to an answer.

**Tip 2: Make the Investment to Measure.** It is easy to collect informal observations and team member input and assume you have identified lessons learned and directions for the future. However, such pictures can be skewed and costly in the long run. The development of quality performance measurement systems requires thought and preparation up front and should be looked at as an organizational investment with the potential to yield great returns. For example, it has been argued that good team performance measurement systems: (a) are based on theory, (b) define measurement purpose(s) up front, (c) utilize team and cognitive task analyses to determine measurement criteria and how to measure them, (d) use a multilevel focus, (e) are practical to use, and (f) involve the user (Dwyer & Salas, 2000). Despite the investment that organizations must make in order to develop quality team performance measurement systems, they also offer significant organizational benefits. For example, making the investment of a multimethod measurement approach that assesses both team process and outcome as well as all relevant competencies (e.g., behavioral, cognitive, and attitudinal) can provide the organization with: (a) a diagnostic picture of team performance, (b) specific and valid directions for improvement, (c) performance data so that trends can be observed (Smith-Jentsch et al., 1998), (d) enhanced employee satisfaction in that their input will contribute to positive changes, and (e) training that can be more efficient in targeted areas in need of remediation.

**Tip 3: Define Measurement Goals.** Once organizations have made decided to make the investment to measure, measurement goals and objective should be defined prior to the design of the performance measurement system. Defining the purpose of measurement is akin to a good theory—it drives the rest of the measurement process. Measurement goals define not only what is measured, but also the level of measurement and form of measurement. For example, measurement goals will determine whether taskwork skills, teamwork skills, or both should be measured. Taskwork skills refer to those skills that team members must understand and acquire for task performance, whereas teamwork skills are those that reflect the behavioral and attitudinal responses that members need in order to function effectively as part of a team (Morgan, Glickman, Woodard, Blaiwes, & Salas, 1986). For example, if the measure-
ment goal is to assess technical training, that would seem to argue for the measurement of taskwork as opposed to teamwork skills.

In addition, depending on the purpose of measurement, assessment protocols may differ (Dwyer & Salas, 2000). For example, if the intent is to examine overall performance or trends across time, outcome measures may be appropriate. Conversely, if the goal is diagnosis, then process measures need to be developed and implemented (Oser, Cannon-Bowers, Dwyer, & Salas, 1997). Taking this a step further, Cannon-Bowers and Salas (1997) have developed a framework that begins to match up specific types of measurement tools via measurement level (team, individual) and type of measurement (e.g. process, outcome)—see Tables 10.5 and 10.6. The authors point out that they are not suggesting that people sample from only one quadrant, but that the nature of the task and nature of the competencies measured will drive the choice of the most appropriate measurement tool for a given situation. For example, when performance is highly dynamic, online observational scales may be necessary to capture performance because they are unobtrusive and do not pose additional workload on trainees. Under more static conditions it may be possible to conduct protocol analysis; if a high degree of behavioral discretion is required then expert raters may be needed to rate (Cannon-Bowers & Salas, 1997).

**Tip 4: Use Multimethod Approach to Measurement.** In order to capture the dynamic nature of teamwork, measurement systems need to use a multimethod approach to measurement. There are at least two reasons for this. First, teams are composed of individuals and are nested within an organizational context, so measurement approaches should acknowledge this. Specifi-

| TABLE 10.5 |
| Framework for Developing Team Performance Measures |

<table>
<thead>
<tr>
<th>Process</th>
<th>Team</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared task models</td>
<td>Assertiveness</td>
<td>Information exchange</td>
</tr>
<tr>
<td>Cue-strategy association</td>
<td>Task-specific role responsibilities</td>
<td>Procedures for task accomplishment</td>
</tr>
<tr>
<td>Task organization</td>
<td></td>
<td>Task strategy association</td>
</tr>
<tr>
<td>Compensatory behavior</td>
<td></td>
<td>Mutual performance monitoring</td>
</tr>
<tr>
<td>Collective efficacy</td>
<td></td>
<td>Flexibility</td>
</tr>
<tr>
<td>Dynamic reallocation of functions</td>
<td></td>
<td>Accuracy</td>
</tr>
<tr>
<td>Task interaction</td>
<td></td>
<td>Latency</td>
</tr>
<tr>
<td>Mission/goal accomplishment</td>
<td></td>
<td>Errors</td>
</tr>
<tr>
<td>Aggregate latency</td>
<td></td>
<td>Safety</td>
</tr>
<tr>
<td>Error propagation</td>
<td></td>
<td>Timeliness</td>
</tr>
<tr>
<td>Aggregate accuracy</td>
<td></td>
<td>Decision biases</td>
</tr>
</tbody>
</table>

TABLE 10.6
Measurement Tools Useful to Assess Team Performance in Training

<table>
<thead>
<tr>
<th>Team</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>Observational scales</td>
</tr>
<tr>
<td></td>
<td>Expert ratings</td>
</tr>
<tr>
<td></td>
<td>Content analysis</td>
</tr>
<tr>
<td></td>
<td>Protocol analysis</td>
</tr>
<tr>
<td>Outcome</td>
<td>Observational scales</td>
</tr>
<tr>
<td></td>
<td>Expert ratings</td>
</tr>
<tr>
<td></td>
<td>Critical incidents</td>
</tr>
<tr>
<td></td>
<td>Automated performance recording</td>
</tr>
</tbody>
</table>


cally, team performance measurement needs to exist at multiple levels. At the simplest level there are individual competencies that are held with regard to the task and the team, and there are team-level competencies held at the individual level with regard to the task and team. Measurement tools should be developed that capture information at both levels. In addition, organizational-level factors may impact the team’s ability to perform, and when possible, measurement systems should take these into account in order to correctly assess remediation strategies for identified deficiencies. Unlike teams created in a laboratory, teams operating within naturalistic environments do not operate in a vacuum.

Different types of measurement tools may be needed to assess the various competencies depending on not only the type of competency being assessed, but also the level at which it is being assessed. Furthermore, the utilization of multiple measurement techniques provides a more complete picture in that each technique has strengths and weaknesses. Multiple measurement techniques may also help combat problems with same method bias.

Tip 5: Focus on Process as Well as Outcome. Team performance measurement within organizations should include the collection of both process and outcome data. Though outcome data is typically collected within organizations due to its focus on the bottom line, all it tells the organization is how the team is performing overall. Although this provides an important piece of information to organizations, it does not provide any information that enables diagnosis, and if collected in isolation, can present a false picture. For example, the team could have achieved an acceptable outcome due to factors beyond their control or by a stroke of luck. In addition, if members are to work together again in the future, there are other indices that need to be taken into account—the interaction processes.

In order to diagnosis a team’s strengths and weaknesses and suggest areas for remediation process, measures must be collected also. Team process variables
that seem to be especially relevant to the organizational teams of the 21st century are those skills that are transportable such as: backup/monitoring behavior, adaptability, communication, decision making, coordination, self-correction behaviors, and situation awareness. It is also important to realize that both process and outcome are needed in order to get a complete picture of team performance because collecting only process behavior can lead to misleading information. For example, what about the case where the organization trained the wrong team and taskwork competencies? The team process might be great, but the resulting impact on the organization's bottom line (e.g., output) may be negligible.

Tip 6: Measure Long-Term Changes. Though it is tempting to measure only short-term outcomes of either a training intervention or team performance as they are the easiest to measure, it is also important to measure more distal outcomes. There are several reasons why both types of measurement are needed. First, due to the dynamic nature of teamwork, a single snapshot of team performance will likely be insufficient, especially if taken early in team development (Morgan et al., 1986). Team performance needs to be sampled over a wide variety of conditions and times to get an accurate picture. Related to this is that many interventions result in temporary improvements that disappear as the newness wears off. Furthermore, depending on the cycle of team development, interventions may have differential impact. For example, Bowers, Barnett, Weaver, and Stout (1998) argue that changes may occur in shared knowledge and the requisite communication behaviors depending on team member experience. Early on, members of effective teams are expected to have more explicit communication due to shared knowledge structures still being formed; later on, communication becomes more implicit in effective teams.

CONCLUDING COMMENTS

There is no question that work teams are here to stay and are often viewed as organizational Band-Aids. However, history has shown us that work teams are not always effective. Furthermore, organizational work teams operating in the 21st century are finding that technology and global competition are changing the venues within which they operate. As a result team members often operate under conditions of role ambiguity where members (a) have distributed expertise, (b) are required to deal with more information in a shorter time frame than ever before, (c) may be culturally diverse, (d) may be nonhuman, (e) are distributed in space and time, and (f) pay a high consequence for errors. Within this environment, the systematic evaluation of team performance becomes a paramount concern as it is the only method by which team members can receive reliable, timely feedback that allows them to correct ineffective cog-
nitions, behaviors, and attitudes that exist within the team—thereby allowing the team to avoid costly errors and consistently meet performance goals.

As researchers we have learned much about the development and implementation of sound, reliable team performance measurement systems, but have not done a good job of marketing what we have learned to those “in the trenches.” Ironically these are the very people who most need this knowledge. Within the current chapter, we sought to take a step in this direction by translating what we know about team performance measurement and enumerating challenges that those in the trenches might face when trying to implement our advice. We also sought to identify what was being done to meet some of the challenges and extracted some practical tips for practitioners charged with team performance measurement. In doing so we found that much progress has been made in trying to meet measurement challenges, but there still remains much to be done. The two largest challenges that face those in the trenches may be determining what to measure and then developing instruments that are reliable, provide diagnostic information, yet are also practical. Finding the correct balance between sound measurement practices and practicality of implementation and scoring will remain a key challenge for practitioners and scientists alike. Though our measurement systems will never be perfect, the team performance measurement community must continually seek to push the envelope.

REFERENCES


Fowkes, J. E., & Burke, C. S. (2005). Targeted acceptable responses to generated events or tasks (TARGETs). In N. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.), Handbook of human factors and ergonomics methods (pp. 53.1–53.6). Boca Raton, FL: CRC Press.


Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of team mental models: We have no shared schema. Organizational Research Methods, 3(2), 123–165.


