

# Chapter 2

## The Assessment of Team Performance: Observations and Needs

Eduardo Salas, Denise L. Reyes, and Amanda L. Woods

**Abstract** The abundance of teams within organizations illustrates the importance of team performance measurement—tools that measure teamwork. Taking into account the inherently complex nature of teams, this chapter presents a few insights and a picture of the research and practice on teamwork measurement over time. We define what makes a team and identify the characteristics of an effective team. Then, we present critical observations to team performance measurement that reflect the 30 years of experience of the first author, at observing, measuring, and assessing team performance in various domains. These observations provide insight into what attitudes, behaviors, and cognitions—how teams feel, act, and think—play an integral role in performance assessment, while taking situational factors and construct considerations into account. Support is presented from the literature on teams and performance measurement, and we provide major contributions from a sample of team performance measurement literature in the past 30 years. We conclude with a discussion on needs for developing future team-based measurement approaches. In this discussion of the future, emphasis is placed on our need, as a field, to continue closing the gap between research and practice through designing and validating effective performance-based measures that target practitioner needs.

**Keywords** Teamwork · Performance measurement · Assessment

---

E. Salas (✉) · D.L. Reyes · A.L. Woods  
Rice University, Houston, USA  
e-mail: Eduardo.Salas@rice.edu

D.L. Reyes  
e-mail: Denise.L.Reyes@rice.edu

A.L. Woods  
e-mail: Amanda.L.Woods@rice.edu

## 2.1 Introduction

Teams are a way of life in organizations. The military, the aviation and space industry, healthcare, corporations, and educational institutions all depend and rely on teams today more than ever. Effective teamwork creates knowledge, minimizes errors, promotes innovation, saves lives, enhances productivity, increases job satisfaction, and ensures success. Teams, when deployed, trained, and led correctly, can be powerful. But insuring that teams perform, learn, develop, and mature is not easy. In fact, it is complex and difficult. A key component to help with this is performance measurement—tools that measure teamwork. Thus we need to create these tools to accurately determine the strengths and weaknesses of the team. This is not an easy goal. We need valid, reliable, theory-driven practices that account for the dynamic nature of teams (Brannick & Prince, 1997; Langan-Fox, Wirth, Code, Langfield-Smith, & Wirth, 2001). This is a tall order, but progress has been made—much progress; this volume is a testament of that progress.

This chapter contributes to the volume by presenting a few insights and a picture of the research and practice on measuring teamwork over time. We will first provide some definitions to set the stage. We will next present some critical observations about measuring team performance. These observations are based primarily on the 30 years of experience of the first author at observing, measuring, and assessing team performance in various domains. We also rely on the literature to support these observations. Lastly, we will discuss some needs for developing future team-based measurement approaches.

## 2.2 Some Definitions

A team consists of two or more people who have defined roles and depend on each other to accomplish a shared goal (Salas, Dickinson, Converse, & Tannenbaum, 1992). In order to understand how teams work and subsequently perform, we have to understand how much the team knows, what skills they possess, and the overall attitude that they bring to the table; we refer to these elements as *team competencies* (Rosen et al., 2008).

The nature of teams is inherently complex, because individual workers are nested in teams, which are nested in organizations (Cannon-Bowers & Salas, 1997; Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995). With teams adding this dynamic layer of complexity, it is critical to slice apart and analyze what characteristics are embedded in the team, as well as the various factors (e.g., individual, team, and organizational factors) that contribute to team performance (Marks, Mathieu, & Zaccaro, 2001). The first step to understand team performance is to identify what characteristics the team possesses starting out. Examples of these *inputs* are individual motivation, attitudes, and personality traits (Driskell, Salas, & Hughes, 2010). Team-level inputs include power distribution, cohesion, and team

resources (Marks et al., 2001). However, inputs are not limited to these characteristics. The type of task and how complicated it is also play a role. Next, we have to identify the *processes*, or the actions that occur when the team is working together to complete a task (LePine, Piccolo, Jackson, Mathieu, & Saul, 2008; Marks et al., 2001). Thus, it is apparent that teams are riddled with complexity, even at their nascent stages.

Though assessing team performance is challenging, we do it because team performance is linked to team effectiveness. Salas, Stagl, Burke, & Goodwin, (2007) defined team effectiveness as the result of a judgment process whereby an output is compared to a subjective or objective standard. Essentially, the results of the team's inputs and processes are evaluated. Therefore, to ensure accuracy, we must match the outcome with the correct methods of measurement (Rosen, Wildman, Salas, & Rayne, 2012). The team yields outcomes at the team and individual levels. Team-level outcomes require the effort of all team members, such as coordination and communication. Individual-level outcomes include a team member's attitude toward the team, which is related to team performance. Organizational-level outcomes are the resulting products of the task and how the team impacts the overall organization. Before we move on, it is important to remember that individual changes in attitude, motivation, mental models, and task knowledge, skills, and attitudes (KSAs) can impact future team processes and performance outcomes, because individuals make up a team (Cannon-Bowers et al., 1995; Tannenbaum, Beard, & Salas, 1992). Taking all these factors into consideration, in order for us to improve performance assessment, we must adopt a multilevel approach (individual, team, and organizational) to understand all the elements contributing to the way team members work together and what they produce based on their actions. With all of these issues in mind, we will now present our observations (in no particular order).

## 2.3 Observations

### 2.3.1 *Observation 1: We Know a Lot*

Team performance measurement is not a perfect science, yet. However, we have learned a great deal over the past 30 years, and we have amassed a robust body of literature on this area of measurement in an effort to address issues that researchers and practitioners face (Brannick & Prince, 1997; Cooke, Kiekel, & Helm, 2001; Kozlowski & Bell, 2003; Rosen et al., 2012; Wildman et al., 2012). Rosen and colleagues (2013) elucidated key components of team performance, as well as providing helpful guidelines for assessment in the context of performance in healthcare settings. Kendall and Salas (2004) addressed methodological concerns by investigating reliability and validity issues impacting team performance metrics. Taking a finer lens to team processes, He, von Davier, Greiff, Steinhauer,

and Borysewicz (2015) have made significant progress towards the development of assessments (e.g., the Programme for International Student Assessment [PISA]) that capitalize on current technology to capture team collaborative problem-solving abilities. Due to recent research efforts, the ability to objectively capture real-time performance is also on the horizon (Stevens, Galloway, Lamb, Steed, & Lamb, 2017). To summarize, we know about why, how, when, and what to measure, but gaps remain. We will talk more on this later; for a more in-depth glimpse into team performance measurement advances, refer to Table 2.1.

**Table 2.1** Sample of team performance measurement literature in the past 30 years

Source	Major contribution(s)
Kendall and Salas (2004)	Examined the criterion problem of team performance, explained current methods for measuring teamwork, and addressed issues of reliability and validity of the measures
Cooke et al. (2000)	Reviewed methods for measuring team knowledge (cognition), a component of teamwork skills and provided methodological needs for the measurement
Salas, Priest, and Burke (2005)	Discussed perceived challenges for those who are responsible for the development of team performance measurement systems, which include defining the purpose, selecting suitable scenarios to use, accounting for timing of the measurement, quantifying responses of teams, and determining how to simplify the collection of data
Salas, Burke, Fowlkes, and Priest (2004)	Explained research in a style for organizations to understand more about the basic elements of team performance measurements, practical requirements for evaluating teamwork skills, and tools to evaluate team skills in order to implement measurements in applied settings
Fowlkes et al. (1994)	Developed Targeted Acceptable Responses to Generated Events or Tasks (TARGETS), an event-based approach for measuring behaviors in teams
Rosen et al. (2013)	Defined key elements of team performance. Provided a guide for measuring, assessing, and diagnosing team performance in healthcare systems
Rosen et al. (2012)	Addressed the challenges faced in measuring team dynamics in real-world settings. Defined team performance measurement methods are defined and presented best practices for developing practical measurements
Salas, Burke, and Fowlkes (2005)	Provided a brief overview of team performance measurement over the past 20 years. Provided a taxonomy of teams present in organizations along with the challenges of measuring their performance. Discussed how these challenges are currently being addressed and offered practical suggestions for practitioners

### ***2.3.2 Observation 2: Context and Purpose of Measurement Matter***

There is no “silver bullet” when creating a team performance measurement tool. We need to think about the context when creating all aspects of a measurement; who, how, and what is being used to conduct the evaluation. Team size, complexity of the task, physical environment of the task, task interdependence, and the amount of communication and interaction required to complete the task should also be considered (Salas, Burke, & Fowlkes, 2005).

The purpose of the performance measurement (i.e., team feedback) should determine what will be collected, and what needs to be collected should determine what kinds of resources are being used for the measurement (Meister, 1985). When choosing a team performance measurement, it is important to remember that all measures need adjustments and modifications in order to have a suitable quality for the required purpose (Salas et al., 2015). Targeting the idiosyncrasies within the team will give you a better idea of what modifications need to be made.

### ***2.3.3 Observation 3: It Is Best to Triangulate***

When it comes to measuring teamwork, it is nearly impossible to collect all of the necessary data from just one source. As noted by Dickinson and McIntyre (1997), “it surely takes a group or team of observers to obtain the necessary information to measure all instances of teamwork” (p. 37). There are a number of ways in which data can be collected. One can use self-report, peer assessments, observations, and objective outcomes. Using different types of data collection is optimal for getting the most data. It is best to use a combination of both qualitative and quantitative data. Subjective ratings are subject to bias; however, there are ways to reduce this bias. For example, observer ratings need to involve interrater reliability to make sure that the variable is being rated accurately from the beginning to the end (Rosen et al., 2012). We can do this by randomly selecting sessions for more than one rater to code and then comparing their ratings (Shrout & Fleiss, 1979). Also, different raters can focus on different areas based on their expertise. For example, supervisors can be used for summative assessments, while peers or subordinates can rate for ongoing or developmental evaluations.

Since teamwork is performed by individuals, it is also important to measure team performance at the individual level. We can achieve a more accurate evaluation of team performance when it is measured at multiple levels. Analysis at the individual level can pinpoint the members who effectively demonstrate teamwork skills (e.g., leadership, coordination, communication). Also, measuring both processes and outcomes can extend the amount of information you can learn about the team’s performance. Looking at processes can give you diagnostic information that addresses issues of development and can serve as a guide for feedback.

Outcome measures, on the other hand, can provide you with “bottom line” performance. Making sure that you have a triangulation approach to collecting data can help ensure validity and address the limitations of the approaches when they are used alone. You do not want any potentially useful data to go unnoticed!

### ***2.3.4 Observation 4: Team Size Matters***

Teams come in all shapes and sizes. When it comes to performance, the size of the team can actually make a difference (Dyer, 1984; Sundstrom, De Meuse, & Futrell, 1990). Hackman (1987) suggested having teams with the least amount of people that are necessary to perform the task. The more team members that are added to a group, the lower the cohesion (McGrath, 1984) as well as group performance (Nieva, Fleishman, & Reick, 1978). The size of a team can be determined by the task at hand or the type of team (i.e., human-computer, distributed teams).

Larger teams run into issues of less flexibility and more differences within the team. More people mean more individual differences. These challenges also carry into the way the team’s performance is measured. Team performance measurement for large teams should include contingency planning, implicit coordination during task execution (i.e., shared mental models), information management, developed understanding of subteams, and an assessment of intra- and interteam cooperation. When conducting observations in complex team settings, raters should not observe more than two team members. This helps to avoid overlooking interactions (Dickinson & McIntyre, 1997).

### ***2.3.5 Observation 5: Subject Matter Experts Can Assess Only Four or Five Constructs***

Experts cannot assess or distinguish more than five team-based constructs. Measuring a construct requires subject matter experts (SMEs), who are individuals that have a strong understanding of the task setting and must make judgments about different team-based constructs. There is a tendency for observers and practitioners alike, to measure all they can measure—sometimes 12 to 14 constructs! Again, raters cannot distinguish these constructs; they all correlate at the end. Our experience is that raters should be trained to focus on only four or five constructs to avoid redundancy (Smith-Jentsch, Zeisig, Acton, & McPherson, 1998). When more than five related constructs are examined, the dimensions start to overlap and become more correlated with each other, making practical distinctions among teams almost impossible. In this case, less is better. Therefore it is wise to select team-based constructs carefully and use only those that matter for team performance.

### **2.3.6 Observation 6: It Is Best to Capture the ABCs—Attitudes, Behaviors, and Cognitions**

It is best to capture representative attitudes, behaviors, and cognitions of teamwork. Teamwork has all of these elements. Noting Observation 5 above, it is best to choose one or two relevant ABCs to capture. Fortunately, an extensive body of research exists surrounding essential ABCs that promote effectiveness. This provides a clear outline of what measurement should capture. Recently, team orientation has been identified as a core attitudinal component of high performing teams (Salas, Sims, & Burke, 2005). Effective teams also promote a wide variety of behaviors such as communication, coordination, and cooperation, to name a few (Campion, Medsker, & Higgs, 1993; Kozlowski & Bell, 2003). For a more in-depth look at team behaviors, refer to Rousseau, Aubé, and Savoie (2006).

Regarding team cognition, shared mental models play an important role in ensuring that team members are on the same page. Successful development of shared mental models helps aggregate the knowledge of each member on the team to create a common understanding of what, how, and when the team needs to accomplish a goal or task (Mathieu, Heffner, Goodwin, Salas, & Cannon-Bowers, 2000). For further discussion of team cognition and its component parts, refer to DeChurch and Mesmer-Magnus (2010). Taken as whole, capturing ABCs is critical for determining how to best measure a team and maximize performance outcomes.

Capturing attitude is commonly used for measuring team performance because it is easy and does not rely on many resources. Measuring attitude is as simple as having team members individually answer a set of items, using a Likert scale to express their feelings in regard to particular statements. Recently, we have also seen examples of attaining information signals by capturing facial expressions, gestures, posture, and periods of silence (Anders, Heinzl, Ethofer, & Haynes, 2011; Shippers, Roebroek, Renken, Nanetti, & Keysers, 2010; Schokley, Santana, & Fowler, 2003; Stevens et al., 2017). We need to measure attitudes because they are associated with team performance (Hackman, 1990; Peterson, Mitchell, Thompson, & Burr, 2000). In regard to behaviors, these can easily be captured through observation. We will elaborate more on what behaviors need to be observed in Observation 7. As for team cognition (knowledge), it still remains a challenge to find a promising method to measure this construct, but it is important to measure because it affects performance (Liu, Hao, von Davier, Kyllonen, & Zapata-Rivera, 2015). In a methodological review, Cooke, Salas, Cannon-Bowers, and Stout (2000) explained that we need to go beyond typical assessments to understand the structure of team knowledge. Different aspects of measurement for this construct include elicitation method (e.g., self-report, eye tracking, communication analysis), team metric, and aggregation method. Nonetheless there is a lot more to be done in regard to measuring cognition (e.g., Wildman et al., 2012).

### ***2.3.7 Observation 7: Behavioral Markers Matter***

Behavioral markers are paramount in performance measurement (Flin & Martin, 2001). Accurately capturing observable behaviors within a team is critical to assessing a team's attributes. These markers should be studied in the context of the environment in which they are being applied. However, mapping constructs to the environment is only a part of the battle. Behavioral markers must be specific and the constructs of interest need to be clearly defined. We have already touched upon various widely used measurement tools that rely on observable team behavior in Observation 6, but a more granular lens must be used to establish what behaviors are of interest. To accurately execute this, time should be taken to methodically carry out the subsequent steps. First, we must establish the behaviors of interest. Next, we must systematically map constructs onto the behaviors. Additionally, we must clearly define the identified constructs. Finally, we must contextualize the behavioral markers by assessing them in the actual performance environment.

### ***2.3.8 Observation 8: It's All About the Constructs, Not the Method!***

A primary issue surrounding constructs is the heightened emphasis placed on the method at the expense of unique traits present in the team. It is important to remember that all teams are not equal! Teams possess both explicit (e.g., observable behaviors such as verbal communication) and implicit qualities (e.g., unobservable processes such as shared mental models; Entin & Serfaty, 1999; Rosen et al., 2012). Due to the developmental nature of teams, certain phenomena (e.g., implicit qualities) emerge in teamwork that can be difficult to capture. Research has attempted to overcome this challenge by placing primary emphasis on the tools used to assess teamwork, but this can sacrifice important aspects of teamwork that influence performance. Most available tools are limited to assessing observable behaviors, but some of the team's most important interactions are implicit and therefore difficult to capture. To illustrate this, in an operating room a patient goes into cardiac arrest; a nurse immediately hands the surgeon necessary tools while the anesthesiologist monitors the patient's current condition and the surgeon attempts to stabilize the patient. This is a good example of a scenario in which implicit coordination is key to the success of the surgical team. Many of these actions need to take place in a matter of seconds; the actions are highly interdependent and do not require explicit communication. As you can imagine, it would be difficult to measure how aligned the team's shared mental model was or how this impacted their ability to coordinate in a highly stressful situation.

Another challenge that centers on the constructs involved in measurement is the statistical method used in analysis. Though accurate and appropriate statistical analysis is critical to team assessment, it does not sufficiently capture performance



all by itself. Many methods of analysis exist that establish the reliability and validity of constructs, but researchers should proceed with caution so as not to become completely reliant on these analyses. The environment and situation being assessed should also play a critical role, to ensure that empirical constructs translate to practical settings (Rosen et al., 2012).

Taking these factors into account when defining constructs is crucial to developing accurate and adaptable performance measures specific to the team. When it comes to teams, adaptability is key (Rosen et al., 2013) and should be reflected in the measurement process. Contextualization should, again, be taken into account, aligning constructs with team competencies to provide accurate construct definitions (Cannon-Bowers et al., 1995).

### ***2.3.9 Observation 9: Measurement of Teamwork Is Not a “One-Stop Shop” Dynamic Phenomenon***

Adding to the complexity of teamwork is the simultaneous need for multiple measurement methods that address the episodic nature of team processes. Teams do not run on fixed intervals; they accomplish different tasks at different times. Hence, it is important to recognize that there is no universal form of measurement that captures performance (Rosen et al., 2012), but keen observation can be a powerful tool when selecting a form of assessment (Rosen et al., 2012).

Although it might be a labor-intensive process to obtain these data, there are new unobtrusive approaches that are promising for team performance measurements. The most popular approaches for observing behavior are event-based measurement, real-time assessment, classification schemes, coding, and behavioral rating scales.

Event-based measurement plays out a scenario where the training objectives are connected to what exactly needs to be assessed. This lets the assessor design events specific to the behaviors to be evaluated. Having control over the events enhances the measurement reliability. Two measurement tools that were developed using the event-based approach are targeted acceptable responses to generated events or tasks (TARGETS; Fowlkes, Lane, Salas, Franz, & Oser, 1994) and team dimensional training (TDT; Smith-Jentsch et al., 1998). One of the most common approaches uses behavioral rating scales such as the behaviorally anchored rating scales (BARS), introduced by Smith and Kendall (1963). Other rating scales include behavioral observation scales (BOS) and graphic rating scales (Latham & Wexley, 1977; Patterson, 1922).

For capturing performance, assessment tools should take on a multilevel perspective (e.g., individual, team, and organizational levels), to accommodate the changes that teams encounter through their life cycle (Rosen et al., 2012; Wildman et al., 2012). Performance should also be measured frequently through a variety of techniques to prevent method bias. However, a challenge this poses is the overuse

of dimensions or measures. Frequently assessing a team can get in the way of team dynamics or otherwise alter the team's normative behavior.

Unobtrusive measures are useful in situations where a team's performance is constantly changing, because they do not disrupt the workflow of the team members. The electroencephalography (EEG) approach to capturing team performance has also shown promise in regard to being unobtrusive while allowing for the real-time assessment of behaviors (Stevens et al., 2017). Automated performance measures (e.g., sociometric badges and audio recording devices) have also shown promise with regard to being both unbiased and unobtrusive. Expounding further on the area of automation, PISA made strides towards capturing both cognitive and social aspects of collaborative problem solving through computer-based assessment (He et al., 2015). One caveat about this method is that automated performance measures are not "stand alone" measures. They still need to be coupled with nonautomated forms of measurement. However, this need for multiple measures holds true for many forms of performance assessment.

### ***2.3.10 Observation 10: What Is Good for Science Is Not Necessarily Good for Practice***

Bridging the gap between research and practice is a critical focus for assessing teamwork performance. This is a challenge because what is good for team research is not always what practitioners want. Researchers can assess many elements of teamwork performance in a controlled laboratory setting, but this freedom can cause researchers to lose sight of what is relevant for practice. Practitioners need tools that are unobtrusive, diagnostic, economical, and easy to use (Rosen et al., 2012). Researchers do not always take an approach that meets these needs. This disparity between research and practice is compounded by the inconsistencies that exist within the dimensions of theoretical teamwork models.

### ***2.3.11 Observation 11: Don't Ignore the Basics***

It is important to go back to the basics to ensure good practice. The underlying premise behind successful measurement provides a sound foundation for future research and practice efforts. The basics illustrate the guiding principles, emerging trends, and considerations for team performance measurement. Outlining clear constructs that target the attitudes, behaviors, and cognitions pertinent to teamwork, while factoring in the context of the environment, lays the groundwork for effective performance measurement.

Great strides have been made in the area of teamwork performance measurement. An area that shows great promise in particular has been modeling and

simulation (Fiore, Cuevas, Scielzo, & Salas, 2002; Hao, Liu, von Davier, & Kyllonen, 2015). However, more development is needed to maintain focus as we move forward. Some of the challenges that still remain are determining what to measure, developing reliable instruments that are diagnostic, and ensuring that these instruments can be implemented across the life span of the team, while placing a heavy emphasis on practicality. To ensure that new methods of assessment are grounded in a reliable and valid foundation, we must go “back to the basics.”

## 2.4 The Future

### 2.4.1 *Observation 12: We Need Tools that Capture the ABCs of Teamwork Dynamically in Real Time that Are Pragmatic, Relevant, and Unobtrusive*

This is the holy grail of team measurement. That is the next step in the future. Efforts have been made to reach this goal; see promising work in Table 2.2. Future research should aim at improving the effectiveness of team measurement, such as the work being done by Cooke (2015), who noted that measuring interactions can

**Table 2.2** Overview of observations on team performance measurements

Team performance measure observations		References
Context matters	<ul style="list-style-type: none"> <li>– No perfect protocol, technique, or format exists</li> <li>– All need adjustments and modifications</li> <li>– All teams are not created equal</li> </ul>	Meister (1985) Salas, Priest, and Burke (2005)
Best to triangulate	<ul style="list-style-type: none"> <li>– Use self-report, peer assessments, and observations</li> <li>– It takes a team to evaluate a team</li> <li>– Use multiple angles, facets, and components</li> </ul>	Dickinson and McIntyre (1997) Rosen et al. (2012) Shrout and Fleiss (1979)
Team size matters	<ul style="list-style-type: none"> <li>– Size of the team makes a difference</li> <li>– Team size affects performance and how performance measures are implemented</li> </ul>	Dyer (1984) Sundstrom et al. (1990) Hackman (1987)
SMEs can only assess four or five constructs	<ul style="list-style-type: none"> <li>– The more constructs, the more correlated they are</li> <li>– Rater training helps</li> <li>– Observations help in debriefing</li> </ul>	Smith-Jentsch et al. (1998)

(continued)

**Table 2.2** (continued)

Team performance measure observations		References
Best to capture ABCs	<ul style="list-style-type: none"> <li>– New unobtrusive approaches are promising</li> <li>– Low-level metrics are also promising</li> <li>– Cognitions remain a challenge</li> </ul>	Smith-Jentsch et al. (1998) Peterson et al. (2000) Liu et al. (2015)
Behavioral markers matter	<ul style="list-style-type: none"> <li>– Be specific</li> <li>– Define constructs of interest precisely</li> <li>– Take time and be systematic</li> <li>– Contextualize constructs</li> </ul>	Flin and Martin (2001) Kendall and Salas (2004)
Need to focus on constructs	<ul style="list-style-type: none"> <li>– Discipline to define constructs is lacking</li> <li>– Lots of focus on the statistics technique is necessary but not sufficient</li> <li>– Obsession with methodological tool often comes at the expense of the phenomena</li> </ul>	Cannon-Bowers et al. (1995) Rosen et al. (2012)
Measuring teamwork is a dynamic phenomenon	<ul style="list-style-type: none"> <li>– Teams do different things at different times</li> <li>– Measure often</li> <li>– Unobtrusive measures are needed</li> </ul>	Rosen et al. (2012) Wildman et al. (2012)
What is good for science is not necessarily good for practice	<ul style="list-style-type: none"> <li>– Practitioners need simple, easy to use, relevant, and diagnostic measures</li> <li>– Researchers can sometimes afford to throw in the “kitchen sink”</li> </ul>	Rosen et al. (2012)
Don’t ignore the basics	<ul style="list-style-type: none"> <li>– Guiding principles are often ignored</li> <li>– New emerging approaches are needed</li> <li>– More is needed, so we should go back to basics</li> </ul>	Salas, Priest, and Burke (2005) Morgan, Glickman, Woodard, Blaiwes, and Salas (1986)
We need tools that capture the ABCs of teamwork dynamically in real-time that are pragmatic, relevant, and unobtrusive	<ul style="list-style-type: none"> <li>– Aim at improving the effectiveness of team measurement</li> <li>– More unobtrusive measures are needed</li> <li>– Acknowledge the advancement of technology and increased usage of online tools for assessment</li> </ul>	Awwal et al. (2015) Cooke (2015)

easily be done unobtrusively and that more unobtrusive measures are needed. Research also needs to acknowledge the advancement of technology and increased usage of online tools for assessment (Awwal, Griffin, & Scalise, 2015).

## 2.5 Conclusion

It is evident that team performance measures are important throughout many industries, and since not all teams are created equally, it is important to modify the measurement based on the specific team. When a measurement system is developed, it should address the question: *Why do we measure?* This question requires a clear definition of the purpose of the measurement tool (von Davier & Halpin, 2013). The purpose behind measuring performance is to generate research, provide teams with feedback, develop team training, evaluate performance, and plan for the future.

During the development of measurement tools another question you need to answer is: *What areas of performance should be captured?* As previously described, to accurately assess performance, the team should be measured on multiple dimensions and the conceptual elements of the measure should be clearly defined. This leads into the temporal considerations of performance assessment: *When should we measure?* Teamwork should be assessed midway through the performance cycle as well as after the conclusion of the performance episode. This begs the question: *Where should teamwork performance be measured?* Teamwork should be measured both in the field through the use of unobtrusive measures as well as in a synthetic environment (Rosen et al., 2013). Lastly, the proper method of analysis should be selected: *How should we measure performance?* Teamwork should be captured through self-report measures, observation, simulations, and balanced scorecards (Rosen et al., 2013).

**Acknowledgements** This work was supported in part by contract NNX16AB08G with the National Aeronautics and Space Administration (NASA) and contract NBPf03402 with the National Space Biomedical Research Institute (NSBRI) with Rice University. The views expressed in this work are those of the authors and do not necessarily reflect the organizations with which they are affiliated or their sponsoring institutions or agencies.

## References

- Anders, S., Heinzle, J., Weiskopf, N., Ethofer, T., & Haynes, J. (2011). Flow of affective information between communicating brains. *Neuroimage*, 54, 439–446.
- Awwal, N., Griffin, P., & Scalise, S. (2015). Platforms for delivery of collaborative tasks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: methods and approach* (pp. 105–113). Dordrecht, Netherlands: Springer.

- Brannick, M. T., & Prince, C. W. (1997). An overview of team performance measurement. In M. T. Brannick, E. Salas, & C. W. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 3–18). Mahwah, NJ, USA: Lawrence Erlbaum.
- Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology*, *46*, 823–847.
- Cannon-Bowers, J. A., & Salas, E. (1997). A framework for developing team performance measures in training. In M. T. Brannick, E. Salas, & C. W. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 45–62). Mahwah, NJ, USA: Lawrence Erlbaum.
- Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). New York, NY, USA: Wiley.
- Cooke, N. J. (2015). Team cognition as interaction. *Current directions in psychological science*, *24*(6), 415–419.
- Cooke, N. J., Kiekel, P. A., & Helm, E. E. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics*, *5*(3), 297–315.
- Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring team knowledge. *Human Factors*, *42*(1), 151–173.
- DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, *95*(1), 32–53.
- Dickinson, T. L., & McIntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. W. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 19–43). Mahwah, NJ, USA: Lawrence Erlbaum.
- Driskell, J. E., Salas, E., & Hughes, S. (2010). Collective orientation and team performance: development of an individual differences measure. *Human Factors*, *52*(2), 316–328.
- Dyer, J. L. (1984). Team research and team training: A state of the art review. In F. A. Muckler (Ed.), *Human factors review* (pp. 285–323). Santa Monica, CA, USA: Human Factors Society.
- Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, *41*(2), 312–325.
- Fiore, S. M., Cuevas, H. M., Scielzo, S., & Salas, E. (2002). Training individuals for distributed teams: Problem solving assessment for distributed mission research. *Computers in Human Behavior*, *18*(6), 729–744.
- Flin, R., & Martin, L. (2001). Behavioral markers for crew resource management: A review of current practice. *The International Journal of Aviation Psychology*, *11*(1), 95–118.
- Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. *Military Psychology*, *6*(1), 47–61.
- Hackman, J. R. (1987). The design of work teams. In J. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). New York, NY, USA: Prentice Hall.
- Hackman, J. R. (1990). *Groups that work (and those that don't)*. San Francisco, CA: Jossey-Bass.
- Hao, J., Liu, L., von Davier, A. A., & Kyllonen, P. (2015). Assessing collaborative problem solving with simulation based tasks. *International Society of the Learning Sciences Proceedings*.
- He, Q., von Davier, M., Greiff, S., Steinhauer, E. W., & Borysewicz, P. B. (2015). *Collaborative problem solving measures in the programme for international student assessment (PISA)*.
- Kendall, D. L., & Salas, E. (2004). Measuring team performance: Review of current methods and consideration of future needs. In J. W. Ness, V. Tepe, & D. R. Ritzer (Eds.), *The science and simulation of human performance* (Vol. 5, pp. 307–326). Bingley, UK: Emerald Group Publishing.
- Kozlowski, S. W. J., & Bell, B. S. (2003). Work groups and teams in organizations. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 333–375). London, UK: Wiley.

- Langan-Fox, J., Wirth, A., Code, S., Langfield-Smith, K., & Wirth, A. (2001). Analyzing shared and team mental models. *International Journal of Industrial Ergonomics*, 28(2), 99–112.
- Latham, G. P., & Wexley, K. N. (1977). Behavioral observation scales for performance appraisal purposes. *Personnel Psychology*, 30, 255–268.
- LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, 61(2), 273–307.
- Liu, L., Hao, J., von Davier, A. A., Kyllonen, P., & Zapata-Rivera, D. (2015). A tough nut to crack: Measuring collaborative problem solving. In Y. Rosen & S. Ferrara (Eds.), *Handbook of research on technology tools for real-world skill development* (pp. 344–359). Hershey, PA, USA: Information Science Reference.
- Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26, 356–376.
- Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, 85(2), 273–283.
- McGrath, J. E. (1984). *Groups: Interaction and performance*. Upper Saddle River, NJ, USA: Prentice Hall.
- Meister, D. (1985). *Behavioral analysis and measurement methods*. New York, NY, USA: Wiley.
- Morgan, B. B., Jr., Glickman, A. S., Woodard, E. A., Blaiwes, A. S., & Salas E. (1986). *Measurement of team behaviors in a Navy environment* (Technical Report No. NTSC TR-86-014). Orlando, FL, USA: Naval Training Systems Center.
- Nieva, V. F., Fleishman, E. A., & Reick, A. (1978). *Team dimensions: Their identity, their measurement, and relationships* (Final Tech. Rep., Contract DAH19-78-C-0001). Washington, DC, USA: Advanced Resources Research Center.
- Patterson, D. G. (1922). The Scott Company graphic rating scale. *Journal of Personnel Research*, 1, 361–376.
- Peterson, E., Mitchell, T. R., Thompson, L., & Burr, R. (2000). Collective efficacy and aspects of shared mental models as predictors of performance over time in work groups. *Group Processes & Intergroup Relations*, 3(3), 296–316.
- Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S. ... Birnbach, D. J. (2008). Measuring team performance in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare*, 3(1), 33–41.
- Rosen, M. A., Schiebel, N., Salas, E., Wu, T., Silvestri, S., & King, H. (2013). How can team performance be measured, assessed, and diagnosed. In E. Salas & K. Frush (Eds.), *Improving patient safety through teamwork and team training* (pp. 59–79). New York, NY, USA: Oxford University Press.
- Rosen, M. A., Wildman, J. L., Salas, E., & Rayne, S. (2012). Measuring team dynamics in the wild. In A. Hollingshead & M. S. Poole (Eds.), *Research methods for studying groups: A guide to approaches, tools, and technologies* (pp. 386–417). New York, NY, USA: Taylor & Francis.
- Rousseau, V., Aube, C., & Savoie, A. (2006). Teamwork behaviors: A review and an integration of frameworks. *Small Group Research*, 37(5), 540–570.
- Salas, E., Benishek, L., Coultas, C., Dietz, A., Grossman, R., Lazzara, E., et al. (2015). *Team training essentials: A research-based guide*. New York, NY, USA: Routledge.
- Salas, E., Burke, C. S., & Fowlkes, J. E. (2005). Measuring team performance “in the wild.” Challenges and tips. In W. Bennet Jr., C. E. Lance, & D. J. Woehr (Eds.), *Performance measurement: Current perspectives and future challenges* (pp. 245–272). Mahwah, NJ, USA: Erlbaum.
- Salas, E., Burke, C. S., Fowlkes, J. E., & Priest, H. A. (2004). On measuring teamwork skills. In J. C. Thomas & M. Hersen (Eds.), *Comprehensive handbook of psychological assessment* (Vol. 4, pp. 427–442). Hoboken, NJ, USA: Wiley.
- Salas, E., Dickinson, T., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Norwood, NJ, USA: Ablex.

- Salas, E., Priest, H. A., & Burke, C. S. (2005b). Teamwork and team performance measurement. In J. R. Wilson & N. Corlett (Eds.), *Evaluation of human work* (3rd ed., pp. 793–808). Boca Raton, FL, USA: CRC Press.
- Salas, E., Sims, D. E., & Burke, C. S. (2005c). Is there a “big five” in teamwork? *Small Group Research*, *36*(5), 555–599.
- Salas, E., Stagl, K. C., Burke, C. S., & Goodwin, G. F. (2007). Fostering team effectiveness in organizations: Toward an integrative theoretical framework. *Nebraska Symposium on Motivation*, *52*, 185–243.
- Schippers, M., Roebroek, A., Renken, R., Nanetti, L., & Keysers, C. (2010). Mapping the information flows from one brain to another during gestural communication. *Proceedings of the National Academy of Sciences USA*, *107*, 9388–9393.
- Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 326–332.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428.
- Smith-Jentsch, K. A., Zeisig, R. L., Acton, B., & McPherson, J. A. (1998). Team dimensional training: A strategy for guided team self-correction. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 271–297). Washington, DC, USA: American Psychological Association.
- Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, *47*(2), 149–155.
- Stevens, R., Galloway, T., Lamb, J., Steed, R., & Lamb, C. (2017). Linking team neurodynamic organizations with observational ratings of team performance. In A. A. von Davier, M. Zhu, & P. C. Kyllonen (Eds.), *Innovative assessment of collaboration*. New York, NY, USA: Springer.
- Sundstrom, E., De Meuse, K. P., & Futrell, D. (1990). Work teams: Applications and effectiveness. *American Psychologist*, *45*(2), 120–133.
- Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. *Advances in Psychology*, *82*, 117–153.
- von Davier, A. A., & Halpin, P. F. (2013). *Collaborative problem solving and the assessment of cognitive skills: Psychometric considerations (ETS Research Report No RR-13-41)*. Princeton, NJ, USA: Educational Testing Service.
- Wildman, J. L., Thayer, A. L., Pavlas, D., Salas, E., Stewart, J. E., & Howse, W. R. (2012). Team knowledge research: Emerging trends and critical needs. *Human Factors*, *54*(1), 84–111.