From Wiese, C.W., Shuffler, M.L., Salas, E., 2015. Teamwork and Team Performance Measurement. In: James D. Wright (editor-in-chief), International Encyclopedia of the Social & Behavioral Sciences, 2nd edition, Vol 24. Oxford: Elsevier. pp. 96–103.
ISBN: 9780080970868

Teamwork and Team Performance Measurement

Christopher W Wiese, Institute for Simulation & Training, University of Central Florida, Orlando, FL, USA
Marissa L Shuffler, College of Business & Behavioral Science, Clemson University, Clemson, SC, USA
Eduardo Salas, Institute for Simulation & Training, University of Central Florida, Orlando, FL, USA

Abstract

Teams are essential in order to handle the complex problems faced by the contemporary organization. Consequently, a deeper understanding of what makes a team effective has been an ongoing goal of researchers. However, our understanding of effective teams is capped by the quality of measurement used to assess them. This article first provides theoretical basis for understanding the measurement of teams. Based on this information, we further offer guidelines and recommendations for those seeking measure team dynamics.

Introduction

The ever-increasing demands of the modern day organization are, at times, too much for an individual to handle. Contemporary technologies have increased the speed, the amount, and the variety of information organizations are expected to process. A consequence of this growth has been the increased use of teams as an appropriate means for handling complex problems and issues above and beyond individual capacities. As the research on teams have accumulated over the past 30 years – and the use of teams in organizational settings show no signs of slowing down – this article aims to fortify the assessment of teams in organizational settings by focusing on the measurement aspect of teamwork. Specifically, we highlight the importance of having a complete team measurement system, as a common issue arising during team assessment is the deficient measurement of multiple team components. A complete team measurement system gauges the individual characteristics of the team (e.g., knowledge, skills, abilities, and other characteristics), both taskwork and teamwork processes/emergent states (e.g., communication, collaboration, cohesion, conflict), and team outcomes. This article will briefly review a representative sample of what is known about each of these team measurement system components. First, we define teams and elaborate on these three components. Next, we identify six considerations for those developing/selecting team measurement systems. Third, we feature a representative sample of methods and tools that can be used in such systems. Finally, we end with a few remarks about the next frontier in teamwork and its measurement.

Overview of Team Measurement Systems

For our purposes, we define teams as a distinguishable set of two or more individuals who must perform distinct, complementary, or interdependent tasks en route to accomplishing a shared, common goal (Salas et al., 1992). This particular definition lends itself well to the discussion of team measurement systems. In the broadest of terms, a system of team measurement could reflect any battery of techniques used to observe and measure teams. However, a team measurement system is not complete unless it captures the individual characteristics of teams, team processes and emergent states, and team outcomes. One way to think of such systems is through the model displayed in **Figure 1**. This figure depicts two different levels of specification (i.e., individual level and team level) and three different components (i.e., inputs, mediators, and outcomes). In order to have a complete system, it is necessary to measure components in all six sections of the model. Though the measurement of individual-level factors is important, this article will focus on the measurement of team-level factors. In the following section, we will briefly define and discuss each of these team-level components.

Inputs

A team is made up of a collection of individuals whereby each individual has certain knowledge, skills, abilities, and other characteristics that have the potential to affect the work they do while on that team (Barrick et al., 1998). Research has indicated the importance of different types of knowledge (e.g., task knowledge), skills (e.g., technical skills), abilities (e.g., cognitive ability), and other (e.g., personality) characteristics in terms of their impact on performance at an individual level (Bunderson and Sutcliffe, 2003; Carson et al., 2007; Driskell et al., 2006; Earley and Mosakwoski, 2000). However, it has been suggested that the KSAOs related to team effectiveness may differ than those that are effective at the individual level (Morgeson et al., 2005). For instance, Stevens and Campion, (1994) suggest that rather than focusing on more individual-level KSAOs (e.g., technical skills, personality traits), there are specific interpersonal KSAs (e.g., conflict resolution, collaborative problem solving, and communication KSAs) and self-management KSAs (e.g., goal setting and performance management and planning/task coordination KSAs) that are more applicable to understanding teamwork. Other critical team-level input factors include, but are not limited to, social skills, teamwork knowledge, expertise diversity, task interdependence, and leadership skills (Morgeson et al., 2005; Chen et al., 2004; Burke et al., 2006; Mathieu et al., 2008).

Mediators

Mediators of the input–outcome relationship in regards to teamwork refer to those processes and other variables that serve

International Encyclopedia of the Social & Behavioral Sciences, Second Edition, 2015, 96–103

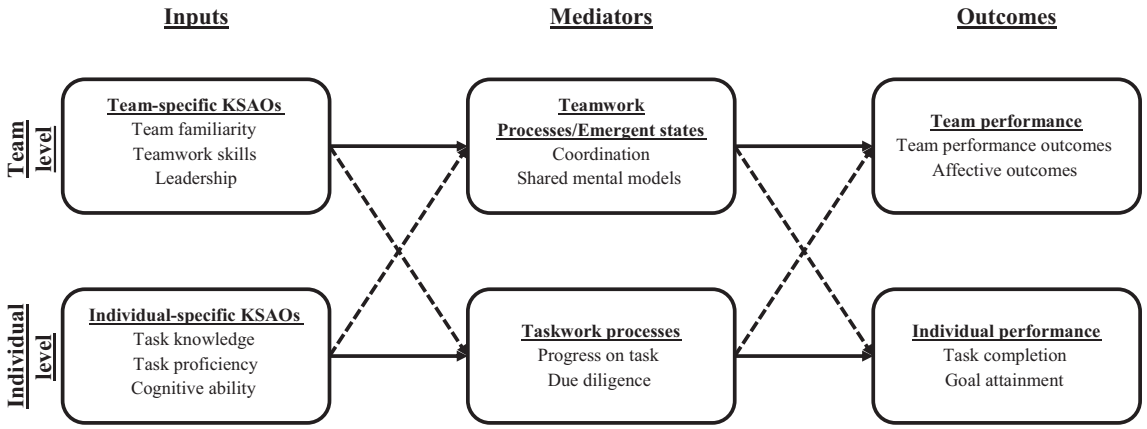

Figure 1 Multi-level team measurement system model. Solid arrows represent a stronger association between components. Dashed lines represent a weaker association between components. Influenced by McGrath, J.E. 1964. Social psychology: A brief introduction, New York, Holt, Rinehart, & Winston; Ilgen, D.R., Hollenbeck, J.R., Johnson, M., Jundt, D., 2005. Teams in organizations: from Input-Process-Output models to IMOI models, Annual Review of Psychology 56, 517–543; and Mathieu, J., Maynard, M., Rapp, T., Gilson, L., 2008. Team effectiveness 1997–2007: a review of recent advancements and a glimpse into the future. Journal of Management 34 (3), 410–476.

as the mechanisms through which individual and team level inputs influence team outcomes (Ilgen et al., 2005). When selecting important team-level mediators, there are two important distinctions to make. First is the difference between taskwork and teamwork. Taskwork is comprised of the actions that team members perform, which are not dependent on any of their fellow team members, whereas teamwork is comprised of those actions made necessary by interdependence and are used to facilitate coordination (Salas et al., 2008). From a measurement standpoint, one could consider actions that represent taskwork to be specific to the individual (e.g., behaviors an individual engages in to complete the team's goal), whereas teamwork needs to be measured at the team level (e.g., interactions between team members). Thus, the focus of team-level measurement is on teamwork mediators.

The second distinction regarding mediators relates to the different types of teamwork. Teamwork can be separated into two different classifications: processes and emergent states. Teamwork processes are the observable behaviors that occur when team members interact with one another, whereas emergent states are cognitive, motivational, and affective states of the team (Marks et al., 2001). Research has demonstrated that both team processes (e.g., strategy formulation and planning, mission analysis, affect management) and emergent states (e.g., team mental models, team cohesion, team empowerment) are crucial predictors of team performance (Edwards et al., 2006; Lepine et al., 2008; Smith-Jentsch et al., 2005).

Outcomes

Team outcomes can be separated into two distinct sets: performance and affective outcomes (Hackman and Morris, 1975). Team performance outcomes are typically denoted by the assessment of the team's accomplishment of assigned goals. The measurement of these outcomes can range from a simple checklist of predefined goals the team was assigned to accomplish to a supervisor's assessment of a team's accuracy and quality of work performed. The assessment of team performance outcomes typically is directly related to the accomplishment of task/team goals. Conversely, affective outcomes target how the team feels regarding their teamwork experience. Some prominent affective outcomes include the team's willingness to work together in the future, team satisfaction, and team member trust. Though some may consider affective outcomes are less important than performance outcomes, they may have critical implications for teams that plan to perform together in the future.

As mentioned earlier, this section served to briefly review some of the more prominent team-level factors crucial in the selection or development of any team measurement system. There are several other factors that are not represented here, but still may be needed in a system depending on the needs of the organization and context of the team. However, the selection of constructs is not the only consideration one should make in developing a team measurement system. In fact, the selection of these constructs is largely influenced by the purpose driving the need for measurement. The next section covers the significant considerations that we believe are necessary in the development or selection of any team measurement system.

Team Measurement System Selection/Development Considerations

Measurement is paramount for the understanding of human behaviors (Meister, 1985). This argument can be extrapolated to all levels of human behavior, an individual as well as a team. Therefore, as the development and use of a team measurement system is vital to understanding intricacies of an effective team, some basic assumptions should be covered regarding the development of such systems (**Table 1**). These assumptions form the basis for the design, development, and testing of team measurement systems. Understanding these underlying assumptions leads to the recognition that there are certain requirements that underlie effective measurement. This section outlines six different considerations that organizations and practitioners should contemplate when developing/selecting a system of measurement for teams. While developing/selecting a team measurement system, it is vital to recognize that

Table 1 Representative sample of basic team measurement assumptions

Basic measurement assumptions
1. The construction of team measures must be based on team-based theoretical models.
2. Team performance and processes are reflected through the actions and observable outcomes of individual team members
3. The goal of measurement is to allow observations of behavior and changes in behavior of team members, which represent underlying processes.
4. Certain team processes and outcomes are more readily available for measurement than others.
5. Processes not easily measured must be matched to corresponding observable behavior or quantified through feedback from team members.
6. Individual performance of team members can be considered within the context of the team and used, in combination with others factors, to assess team performance.

Based on Meister, D., 1985. Behavioral Analysis and Measurement Methods, Wiley-Interscience; Brannick, M.T., Salas, E., Prince, C.W., 1997. Team performance assessment and measurement: theory, methods, and applications, Psychology Press.

these considerations are not independent of one another and therefore should be deliberated in their entirety.

Measurement Purpose

The first consideration facing those who are developing or selecting a team measurement system is to specify the purpose of the measurement. Team measurement systems can be used to serve many purposes, such as team member selection, training feedback, performance evaluation, and promotion. The selection of the purpose will guide measurement content selection, decisions regarding the type of data needed, and appropriate data collection resources (Mesiter, 1985). For instance, if the purpose of the measurement is simply to gauge team members' reactions to training program, a paper-pencil measure specifically targeting their feelings about the training may suffice. However, if the purpose of the measurement is to provide feedback and develop a flight crew, a more labor-intensive, complex system may be necessary. In this latter instance, critical team process constructs would need to be specified, measured, and tied to specific team outcome variables so that detailed feedback could be given regarding the team's performance. Therefore, specifying the purpose of the measure is a critical first step in any team measurement system.

Construct Selection

Once the purpose of the team measurement system is specified, the next step involves selecting the constructs of interest. It is important for developers to measure all three factors previously discussed: team-level inputs, mediators, and outcomes. Incorporating all three elements into a team level system of measurement will provide the most holistic understanding of what factors may be impacting team performance, and how such factors may be interacting with one another. If the decision is made to only select a portion of these factors, the team measurement system may be deficient and unable to accomplish the overall purpose. For instance, only measuring team performance outcomes in a training program would not be conducive to providing feedback to team members, a critical aspect to effective training. In order to provide feedback, specific team processes must be measured. By giving feedback on specific team processes, team members will be able to learn specifically how they were performing improperly. Additionally, the selection of constructs identified as inputs, mediators, and outcomes will also impact when these constructs should be measured, as will further be elaborated upon later in this article.

Measure Referent

When measuring team-level variables, it is imperative that those developing measures understand the implications associated with specifying the focal point or referent of the measure. The focal point or referent of the measure refers to whom or what the measure is attempting to obtain information about; for teams, this could be regarding the entire team as a whole, or it could refer to the individuals of the team (Chan, 1998). This consideration encapsulates several questions: Do external raters need to observe the construct or can the construct be assessed through paper–pencil measures? Are the items on the measure specifically referring to the team instead of the individual? If the items on a measure are referring to the individual, can they simply be aggregated to a team level or is a more complex aggregation technique needed? Is the targeted construct observable? Does the task provide the opportunity for the targeted-construct to be performed? Such questions are critical to ask in order to ensure that the appropriate focal point is utilized for each of the constructs to be measured.

The relevance of these questions is partially determined by the constructs selected. For instance, when assessing team cohesion, it may be necessary to utilize survey items that specifically use the team as referent. Given that team cohesion is an emergent state, which can be very challenging to observe by outsiders (Mathieu et al., 2008), having team members report on their feelings with the team as the referent should provide the most appropriate measure of the construct. Alternatively, if one decides to measure teamwork processes, it may behoove the measurement system designer to choose a situation in which team members would have the opportunity to perform the behaviors associated with the construct prior to measurement in order to ensure that those behaviors have a chance to emerge at the team level. Much like the previous considerations, the decisions made here greatly influences the choices available in the forthcoming steps.

Measurement Source

There are five primary measurement sources typically used in team measurement systems: team members, supervisors, trained raters, customers, and objective sources. Not all of these sources are equally effective in assessing all different factors (i.e., inputs, mediators, and outcomes). For example, the measurement of input factors is more accurate when coming from the team members themselves. Supervisors may not be as accurate at assessing these input factors as they may not be aware of each team members' expertise. A summary of the specific advantages and disadvantages for each source is

Table 2 Pros and cons of measurement sources

Measurement source	*Pros*	*Cons*
Team members	● Observed team members on a daily basis ● Good understanding of what should be done ● Can access their own cognitions	● May be too lenient on team members if team rewards are tied to team behavior ● May be too severe on ratings if the climate is very competitive
Supervisors	● Expert understanding of what constitutes good behavior ● Witness to best behaviors	● May not be around on a daily basis ● Cannot assess team cognitions
Trained raters	● Trained understanding of the behaviors being assessed ● May be experienced in doing ratings	● Not privy to day-to-day operations ● Not familiar with organization/team culture
Customers	● Gives a different perspective compared to the aforementioned sources. ● Ratings come form experience with the team	● Ratings are very heavily influenced by experience ● Does not have an understanding of what constitutes good performance.
Objective sources	● Removes a great deal of bias from team measures. ● Can be concretely defined by the researcher	● Can only measure what it is programmed to measure ● Can typically only measure simple things (e.g., frequency, counts)

provided in Table 2. One must be selective in choosing the source of the measure. As some sources are better at assessing certain factors than others, the inappropriate selection of measurement source could result in inaccurate/biased measurement, which would consequentially lead to improper assumptions made about the team. Thus, it is highly recommended to use multiple sources in collecting team data. Some have even argued that it takes a team to effectively measure a team (Dickinson and McIntyre, 1997).

Measurement Timing

The fact that teams develop and evolve over time has implications on when constructs can and should be measured (Kozlowski and Klein, 2000). Certain constructs may be more appropriate to measure during certain points in a team's performance cycle than others. For instance, teamwork processes targeted at setting up teams for successful performance, such as mission analysis, goal specification, and strategy formulation and planning, may be more observable during the initial stages of team performance. Additionally, the choice of source may have implications on when a construct is measured. For example, externally trained raters can only make judgments regarding targeted constructs while the team is performing. Other team factors should only be measured at the end of team performance (e.g., objective performance, affective outcomes). Additionally, the timing of when constructs are measured may also be influenced by environmental characteristics that are not under the control of the team measurement system designer. If, for example, one of the conditions set forth by a union is that team member can only fill out research surveys when they are clocked in, the developer must try and create a measure that team members can fill out without disrupting the constructs they are trying to assess.

Measure Validation

The last consideration concerns the validation of the measures within a team measurement system. Depending on whether one is selecting or developing the measures in a system, these considerations will change slightly. However, the assurance that these measures have undergone proper validation processes is one of the most crucial steps in the development/selection of a team measurement system. Simply put, validity concerns whether the measure actually captures what it is supposed to measure (Messick, 1980). Though validity can never be absolutely proven, evidence can be provided supporting a measure's validity. There are three main types of validity used to support a measure's credibility: content validity, construct validity, and criterion-related validity. Each of these provides different types of evidence of validity and, consequently, one should look to support all three validity types.

Content validity typically refers to the extent that a measure represents a construct in its entirety. One way to provide evidence for content validity is by using subject matter experts (SMEs) to review a measure for any construct deficiency or contamination. Support for a measure's content validity does not necessarily have to come through an empirical process. That is, support for content validity can be established through a rationalized process of judgments rather than a quantitative investigation. Criterion-related validity simply concerns the degree to which a measure predicts an outcome of interest. Evidence for this type of validity is established when a measure correlates with the criterion. That is, criterion-related validity evidence is typically represented in the form of a correlation coefficient between a measure (predictor) and a criterion. Lastly, construct validity concerns whether or not a measure quantitatively represents the construct that it is designed to represent. There are several ways one can establish construct validity such as creating a nomological network or creating a multi-trait multi-method matrix (Cronbach and Meehl, 1955; Campbell and Fiske, 1959). Each type of validity is critical to the credibility of a team measurement system. As such, providing evidence for each type of validity is a fundamental consideration.

Available Measurement Tools

During the creation of a team measurement system, it is not necessary to 'reinvent the wheel,' or start completely from scratch in terms of developing tools. There are several tools/methods available to assess a team's inputs, mediators, and outcomes. In this section, we cover a representative sample of some of these tools/methods used to measure teams including rating scales, behaviorally anchored rating scales (BARS), event-based measurement scales, automated performance measurement, concept mapping, pathfinder, and probed protocol

analysis. There are numerous cases in recent literature that demonstrate the applicability of these tools for the collection of team data (Jones and Harrison, 1996; Lynn and Reilly, 2000). Moreover, though many of these tools can be used as a combination of inputs, mediators, and outputs, some are more apt at assessing certain constructs than others.

Rating Scales

One of the most common forms of assessment of team constructs is rating scales. Rating scales ask people to indicate their attitudes, opinions, beliefs, or feelings on a large number of items. Thus, rating scales can be used to assess reflections of team behaviors, opinions concerning how team members feel, or even beliefs concerning team member's cognitions. There are a number of ways to develop rating scales. If one desires to gauge opinions concerning team members' behavior, the first step would be to compile a description of the behaviors of interest. One approach to identifying and accumulating these behaviors would be to use a critical incident technique. A critical incident technique gathers information from SMEs regarding memorable events, both positive and negative, which occurred in a specific situation (Chell, 1998). These behaviors would then serve as the content for the questions being asked on the measure.

Once the content has been identified, there are several options on how to present the information. A few possibilities include: (1) scale points can simply be reflected by numbers, (2) descriptive words can be used at end scale points, or (3) verbal descriptions can be used at each scale point (Meister, 1985). Another choice comes with deciding how many scale points to be used. First, there must be enough scale points to fully represent the question being asked as a dichotomous scale might not represent the different degrees of the attribute being assessed. Another consideration is whether to use an odd number of scale points, which provides the rater with the option of choosing a neutral response. These options have their respective pros and cons, and it is completely up to the measure designer which option they prefer. An example of a ratings scale is provided in **Figure 2**.

Behaviorally Anchored Ratings Scales

BARS are a type of rating scale first introduced by Smith and Kendall (1963). Much like rating scales, BARS use verbal descriptors at each scale point. However, unlike rating scales, BARS provide very specific behavioral descriptors at each scale point. These anchors provide more concrete examples of what behaviors represent each scale point. That is, rather than one word representing each scale point (e.g., agree, never, always, disagree), BARS scale points has specific behavioral descriptions. Consequently, this approach lends itself well to the critical incident technique mentioned above. Instead of the behaviors becoming part of the item text (as alluded to above), they are transposed onto the scale points. Furthermore, BARS are especially suitable for assessing team behaviors. One distinct advantage of BARS over traditional rating scales is that it helps control for any individual idiosyncrasies of the raters (Salas et al., 2003). An example is provided in **Figure 3**.

Event-Based Measurement Scales

Another behaviorally oriented approach to team measure is the event-based measurement scales. Using this approach, the measure designer first identifies the team competencies they wish to measure. Following this, they insert events into behavioral scenarios designed to cue the specific behavioral competencies. In doing so, practitioners can tie measurement directly back to the targeted competences. Furthermore, this approach does provide some ease for raters. That is, since the raters know what cues are embedded into the scenario and when these cues take place, they are better able to focus on a certain behavioral competency when the time arises. By creating a situation in which team behavioral ratings are consistently drawn from the same event across teams, event-based measurement scales typically contain exceptional psychometric properties. An example of this type of measurement approach is provided in **Figure 4**.

Automated Performance Measurement

The type of information that automated performance measurement tools can collect is tremendously diverse. Automated performance measurement tools have the ability to collect data as simple as automatically counting the number of times team members say a certain word to data gathered from a highly complex jet flight simulator. One of the advantages of automated performance measurement tools is that they, to a certain degree, remove the errors associated with human decision making. Further, tools like this are not as intrusive as other aforementioned approaches. However, there are several downsides of using an automated performance measurement tool. First, they can be extremely expensive to design and maintain. Next, automated performance measures can only measure what it was programmed to measure. If you want to measure something else, you might not be able to get this information from data that you have already collected. Another

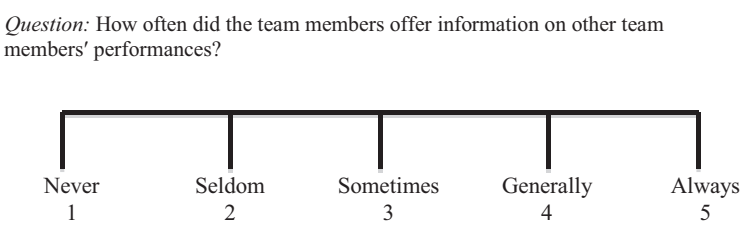

Figure 2 Example of a rating scale measuring feedback.

Figure 3 Example of a behavioral anchored ratings scale measuring feedback.

Flight segment	Event	Target	Hit
Approaching checkpoint	Navigation	Report distance to checkpoint.	
		Report outbound heading from checkpoint.	
At checkpoint	COMMUNICATION from base: SIGMET (SIGnificant METeorlogical information) over checkpoint.	Note implications of SIGMET (e.g., cannot land at checkpoint).	
		Obtain additional information (e.g., from flight service).	
		Make new plan – choose alternates from: • Return to base (best choice) • Pleasanton • Kelly AFB • Randolph AFB	

Figure 4 Events and associated TARGETS for a short-hop segment in which helicopter crew is on a training maneuver (Fowlkes et al., 1994).

downside is that it can only detect overt responses (Meister, 1985). Additionally, these types of tools are not especially well designed at measure team attitudes or cognitions.

Concept Mapping

Concept mapping produces a graphical representation of the knowledge content and structure of an individual (Swan, 1995). Concept mapping has more recently been adapted from the individual level to assess team cognitions. This method creates a map of cognitive concepts that are linked together based on their association with the other concepts. In order to generate the concept maps, the researcher can request the data from team members through interviews/surveys or by post-hoc examination of the data (Mohammed et al., 2000). There are several variations of concept maps, all with their unique purpose. For example, Pathfinder elicits similarity judgments from team members to produce a representation of the team's concept structure (Mohammed et al., 2000). This technique involves the researcher coming up with predefined concepts, usually using

a team task analysis. Participants are then instructed to make pairwise comparison ratings on each concept pairing. After these paired-comparison ratings are made, Pathfinder uses a specialized algorithm to determine the psychological proximity of the concepts as well as a conceptual structure (see Goldsmith and Kraiger, 1997; Cooke and Schvanevedlt, 1988).

Summary

Certain measurement tools are more apt at assessing attitudes (e.g., rating scales), behaviors (e.g., BARS, event-based measurement scales), and cognitions (e.g., concept mapping). When selecting measure to use in a team measurement system, it is recommended that one considers the strengths and weaknesses of each of these measurement tools. Though we emphasize the importance of having a complete team measurement system, if the proper tools are not utilized in appropriate ways, having such a system is meaningless.

The Future of Team Measurement

While much has been discussed thus far regarding the current state of the science of team measurement in terms of important considerations and potential approaches, there is still much to be done. Team measurement as a science is far from being perfect, especially as we continue to explore new and changing team contexts. The following provides a brief summary of two particularly important areas requiring future research, the increasing use of distributed teams communicating via virtual media, and the need to better represent the dynamic and temporal nature of teamwork.

Distributed Virtual Teams

The growing implementation of teams whose members are geographically distributed and communicated virtually is also an issue in need of additional research. Distribution of team members has become commonplace in organizations around the world as a means to deal with complex world problems, and is the focus of a growing body of academic research across multiple disciplines (Connaughton and Shuffler, 2007). Distributed virtual teams are defined as teams whose members are geographically separated and must utilize some type of technology in order to communicate, whether it be something as basic as e-mail or as complicated as a virtual (Driskell et al., 2003). This definition illustrates the major features of such teams: the use of technology to bring together team members that may be spatially distributed (i.e., not colocated), or temporally distributed (i.e., interacting asynchronously; Kozlowksi and Bell, 2003).

Distribution and virtuality are important factors to consider in terms of team measurement as they can change how team members interact with one another. Initial results of both lab and field research illustrate differences in distributed virtual teams when compared to face to face teams (Bos et al., 2006). However, these teams are often investigated using the traditional metric approaches applied to face-to-face teams, and do not take into consideration how distribution and virtuality may be influencing when and how metrics should be captured. Overall, the research regarding how to capture the nuances that may be important to distributed virtual teams is very limited, especially in terms of partial distribution as well as the combined influences of distribution and virtuality. Given the rise toward distributed teams relying upon technology to communicate, it is therefore critical that attention be paid to this issue from a measurement perspective.

Dynamic Assessment of Team Constructs

Although advances in measurement over recent years have led to increasingly effective techniques for capturing team level constructs, additional work is needed in order to effectively capture the dynamic nature of teams. The measurement techniques previously described in this article are effective, but do not always capture the degree to which teams can change rapidly over a period of time. For example, observers may only capture teamwork processes for a short period of time during the performance episode that they are observing, leaving out the richness of how the team may have developed up to that point or after that point. Furthermore, paper-based metrics provided by team members force members to summarize their thoughts and feelings over a period of time, essentially averaging rich data that could be critical to effectively understanding the nature of team performance. Thus, it is critical that future research regarding teamwork and team assessment focus on the dynamic, changing nature of teams over time.

Along these lines, increased attention regarding the implementation of unobtrusive measures in team settings has led to the proposed use of metrics such as physical closeness of team members over a period of time or biometric data (e.g., heart rate, perspiration, eye tracking) as proxies for team level constructs such as trust or cohesion (DeCostanza et al., 2012). Additionally, the need to reduce the reliance on expert observers is also critical to capturing this dynamic nature of teamwork. While metrics such as the BARS and TARGETS techniques described in this article allow less-expert observers to assess team variables, this still incorporates the use of humans in the measurement process, increasing the likelihood of error and mistranslation of information. Furthermore, such observers cannot be present constantly over the lifecycle of a team in every situation; therefore changes that the team may experience in the constructs of interest may not be fully captured. Thus, other ways to reduce this reliance on experts must be developed. In sum, team measurement is a critical concept for both research and practice, as only through effective measurement we will be able to more accurately understand the nuances of teamwork and team performance necessary for enhancing teams and their success in organizations.

See also: Behavior Analysis, Applied; Industrial–Organizational Psychology: Science and Practice; Pay, Compensation, and Performance, Psychology of; Personnel Selection, Psychology of; Program Evaluation.

Bibliography

Barrick, M.R., Stewart, G.L., Neubert, M.J., Mount, M.K., 1998. Relating member ability and personality to work-team processes and team effectiveness. Journal of Applied Psychology 83 (3), 377–391.

Bos, N., Olson, J., Nan, N., Shami, N., Hoch, S., Johnston, E., 2006. Collocation blindness in partially distributed groups: is there a downside to being collocated?. In: Conference on Human Factors in Computing Systems Proceedings, pp. 1313–1321.

Brannick, M.T., Salas, E., Prince, C.W., 1997. Team performance assessment and measurement: theory, methods, and applications. Psychology Press.

Bunderson, J.S., Sutcliffe, J.M., 2003. Management team learning orientation and business unit performance. Journal of Applied Psychology 88 (3), 552–560.

Burke, C., Stagl, K., Klein, C., Goodwin, G., Salas, E., Halpin, S., 2006. What type of leadership behaviors are functional in teams? A meta-analysis. Leadership Quarterly 17 (3), 288–307.

Campbell, D., Fiske, D., 1959. Convergent and discriminant validation by the multitrait–multimethod matrix. Psychological Bulletin 56 (2), 81–105.

Cannon-Bowers, J., Salas, E., 1997. A framework for developing team performance measures in training. In: Team Performance Assessment and Measurement: Theory, Methods, and Applications. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 45–62.

Carson, J.B., Tesluk, P.E., Marrone, J.A., 2007. Shared leadership in teams: an investigation of antecedent conditions and performance. Academy of Management Journal 50 (5), 1217–1234.

Chan, D., 1998. Functional relations among constructs in the same content domain at different levels of analysis: a typology of composition models. Journal of Applied Psychology 83 (2), 234–246.

Chell, E., 1998. Critical incident technique. In: Qualitative Methods and Analysis in Organizational Research: A Practical Guide. Sage Publications Ltd, Thousand Oaks, CA, pp. 51–72.

Chen, G., Donahue, L., Klimoski, R., 2004. Training undergraduates to work in organizational teams. Academy of Management Learning & Education 3 (1), 27–40.

Connaughton, S.L., Shuffler, M., 2007. Multinational multicultural distributed teams: a review and future agenda. Small Group Research 38 (3), 387–412.

Cooke, N.J., Schvaneveldt, R.W., 1988. Effects of computer programming experience on network representations of abstract programming concepts. International Journal of Man–Machine Studies 29 (4), 407–427.

Cooke, N.J., Salas, E., Cannon-Bowers, J.A., Stout, R.J., 2000. Measuring team knowledge. Human Factors 42 (1), 151–173.

Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. Psychological Bulletin 52 (4), 281–302.

DeCostanza, A.H., DiRosa, G.A., Rogers, S.E., Slaughter, A.J., Estrada, A.X., 2012. Researching teams: nothing's going to change our world. Industrial and Organizational Psychology 5 (1), 36–39.

Dickinson, T.L., McIntyre, R.M., 1997. A conceptual framework for teamwork measurement. In: Brannick, M.T., Salas, E., Princem, C.W. (Eds.), Team Performance Assessment and Measurement: Theory, Methods, and Applications. Taylor and Francis, Mahwah, NJ, pp. 19–44.

Driskell, J.E., Radtke, P.H., Salas, E., 2003. Virtual teams: effects of technological mediation on team performance. Group Dynamics: Theory, Research, and Practice 7 (4), 297–323.

Driskell, J.E., Goodwin, G.F., Salas, E., O'Shea, P.G., 2006. What makes a good team player? Personality and team effectiveness. Group Dynamics: Theory Research and Practice 10 (4), 249–271.

Earley, C.P., Mosakowksi, E., 2000. Creating hybrid team cultures: an empirical test of transnational team functioning. Academy of Management Journal 43 (1), 26–49.

Edwards, B., Day, E., Arthur, W., Bell, S., 2006. Relationships among team ability composition, team mental models, and team performance. Journal of Applied Psychology 91 (3), 727–736.

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.

Goldsmith, T., Kraiger, K., 1997. Applications of structural assessment to training evaluation. In: Ford, J.K., Kozlowski, S.W.J., Salas, E., Teachout, M. (Eds.), Improving Training Effectiveness in Organizations. Erlbaum, Mahwah, NJ, pp. 73–96.

Hackman, J.R., Morris, C.G., 1975. Group tasks, group interaction process and group performance effectiveness: a review and partial integration. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, vol. 8. Academic Press, New York, NY, pp. 47–99.

Ilgen, D.R., Hollenbeck, J.R., Johnson, M., Jundt, D., 2005. Teams in organizations: from Input-Process-Output models to IMOI models. Annual Review of Psychology 56, 517–543.

Jones, M.C., Harrison, A.W., 1996. Is project team performance: an empirical assessment. Information & Management 31 (2), 57–65.

Kozlowski, S.W.J., Bell, B.S., 2003. Work groups and teams in organizations. In: Borman, W.C., Ilgen, D.R., Klimoski, R. (Eds.), Comprehensive Handbook of Psychology, Industrial and Organizational Psychology, vol. 12. Wiley, New York, pp. 353–376.

Kozlowski, S.W.J., Klein, K.J., 2000. A multilevel approach to theory and research in organizations: contextual, temporal, and emergent processes. In: Klein, K.J., Kozlowski, S.W.J. (Eds.), Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions. Jossey-Bass, San Francisco, CA, pp. 3–90.

LePine, J.A., Piccolo, R.E., 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.

Lynn, G.S., Reilly, R.R., 2000. Measuring team performance. Research-Technology Management 42 (2), 48–56.

Marks, M., Mathieu, J., Zaccaro, S., 2001. A temporally based framework and taxonomy of team processes. Academy of Management Review 26 (3), 356–376.

Mathieu, J., Maynard, M., Rapp, T., Gilson, L., 2008. Team effectiveness 1997–2007: a review of recent advancements and a glimpse into the future. Journal of Management 34 (3), 410–476.

Meister, D., 1985. Behavioral Analysis and Measurement Methods. Wiley-Interscience.

Messick, S., 1980. Test validity and the ethics of assessment. American Psychologist 35 (11), 1012–1027.

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.

Morgeson, F.P., Reider, M.H., Campion, M.A., 2005. Selecting individuals in team settings: the importance of social skills, personality characteristics, and teamwork knowledge. Personnel Psychology 58 (3), 583–611.

Mortensen, M., Hinds, P.J., 2001. Conflict and shared identity in geographically distributed teams. International Journal of Conflict Management 12 (3), 212–238.

Orasanu, J., 1990. Shared Mental Models and Crew Performance. No. 46. Princeton University, NJ. Laboratory of Cognitive Science, pp. 172–181.

Salas, E., Burke, C.S., Fowlkes, J.E., Priest, H.A., 2003. On measuring teamwork skills. In: Thomas, J.C., Hersen, M. (Eds.), Industrial and Organizational Assessment, vol. 4. John Wiley and Sons, Hoboken, NJ, pp. 427–442.

Salas, E., Cooke, N.J., Rosen, M.A., 2008. On teams, teamwork, and team performance: discoveries and developments. Human Factors: The Journal of the Human Factors and Ergonomics Society 50 (3), 540–547. http://dx.doi.org/10.1518/001872008X288457.

Salas, E., Dickinson, T.L., Converse, S.A., Tannenbaum, S.I., 1992. Toward an understanding of team performance and training. In: Swezey, R.W., Salas, E. (Eds.), Teams: Their Training and Performance. Ablex Publishing, Westport, CT, pp. 3–29.

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.

Smith-Jentsch, K.A., Mathieu, J.E., Kraiger, K., 2005. Investigating linear and interactive effects of shared mental models on safety and efficiency in a field setting. Journal of Applied Psychology 90 (3), 523–535.

Stevens, M.J., Campion, M.A., 1994. The knowledge, skill, and ability, requirements for teamwork: implications for human resource management. Journal of Management 20 (2), 503–530.

Swan, J.A., 1995. Exploring knowledge and cognitions in decisions about technological innovation: mapping managerial cognitions. Human Relations 48 (11), 1241–1270.