Advances in Measuring Team Cognition

Nancy J. Cooke
Arizona State University East
Mesa, AZ

Eduardo Salas
University of Central Florida
Orlando, FL

Preston A. Kiekel
New Mexico State University
Las Cruces, NM

Brian Bell
Arizona State University East
Mesa, AZ

Contact:

Nancy J. Cooke
Applied Psychology Program
Arizona State University, East
7001 E. Williams Field Road
Mesa, AZ 85212

Phone: 480-727-1331
Email: ncooke@asu.edu
Fax: 480-727-1363

Why Measure Team Cognition?

Teams have become an integral and essential social component in organizations. There is no doubt that there are more businesses, industries and agencies implementing team-based systems than ever before. Organizations believe that teams, teamwork and effective team functioning can provide a competitive edge. Teams are, after all, dispatched to tackle difficult and complex problems. Why? Organizations think teams hold the solution to many problems. The perception is that teams can manage stress; teams can adapt and be flexible; team can make better decisions, and teams can be more productive than individuals. So, teams are a popular commodity in organizations.

While the evidence of the efficacy of teams remains open, social scientists have begun to provide much needed answers to this issue. In fact, the amount of research effort and resources aimed at understanding what comprises effective team performance is overwhelming. Military, human factors, social, cognitive and industrial/organizational psychologists have been studying teams for decades (see Levine & Moreland, 1990; Guzzo & Shea, 1992). And so progress has been made. We understand better what teams do, how they do it, why they do it and how we can improve what they do. We can now identify the behaviors (i.e., skills) and attitudes that team need to have in order to be effective (see Salas & Cannon-Bowers, 2000). Moreover, there is evidence that this leads to better performance (see Hackman, 1990; Guzzo & Salas, 1995; Cannon-Bowers & Salas, 1998). But we also know that teams “think”. That is, team members possess knowledge that allows them to function effectively as a team, even during periods of high workload (Orasanu, 1990). This knowledge is as important as skills and attitudes--and in some cases, these cognitive abilities are more important. Team cognition has also been linked to team performance (Stout, Cannon-Bowers, & Salas, 1996). But, unfortunately our understanding of team cognition is limited because it is difficult to capture or measure team cognition. Therefore, to fully understand team performance as well as the contribution of team cognition to effective team functioning, we must measure team cognition, which is the motivation behind this chapter.

Why do we need to measure team cognition? First, measures of team cognition provide a window to some of the factors underlying team acquisition and performance of a complex skill and can thus be valuable in diagnosing team performance successes and failures. Second, with an understanding of team cognition, training and design interventions can target the cognitive underpinnings of team performance. The ability to assess team cognition and predict team performance has far reaching implications for evaluating progress in training programs and diagnosing remediation needs during training. If one can measure the cognitive structures and processes that support task performance, then one can use this information to predict the time course of skill acquisition and the asymptotic levels of performance once skill has been acquired. Finally, understanding the cognition underlying team performance has implications for the design of technological aids to improve team performance not only in training but, more importantly, in actual task environments.

We have organized this chapter into six sections. We first discuss the nature of team cognition. We then highlight some of the limitations of measures followed by a brief discussion of some new approaches to measuring team cognition. Next, we outline some validation procedures and what we have learned so far about team cognition. We conclude with a few remarks about the future of team cognition.

What is Team Cognition?

Teams perform cognitive tasks. That is, they detect and recognize pertinent cues, make decisions, solve problems, remember relevant information, plan, acquire knowledge, and design solutions or products as an integrated unit. But, is team cognition anything more than the sum of the cognition of the individual team members? Yes, our position, which serves as the basis for our approach to measuring team cognition, holds that team cognition is more than the sum of the cognition of the individual team members. Instead, team cognition *emerges from the interplay of the individual cognition of each team member and team process behaviors.*

We see these two intertwined aspects of team cognition, individual cognition of team members and team process behaviors, as analogous to cognitive structures and cognitive processes at the individual level. For example, individual knowledge may be, in principal, described in terms of a knowledge structure, but the effective application of that knowledge structure requires processes such as retrieval and pattern recognition that act on that structure. In a similar way, team members interact through communication, coordination, and other process behaviors and in doing so transform a collection of individuals' knowledge to team knowledge that ultimately guides action. In other words, the outcome of this transformation is effective team cognition. Figure 1 graphically represents this view of team cognition.

[Insert Figure 1 about here]

As is represented in Figure 1, our work to date has focused predominantly on team *knowledge* which we view as central to team *cognition*. Parallel to research on individual expertise (e.g., Chase & Simon, 1973; Glaser & Chi, 1988), accounts of effective team performance highlight the importance of team knowledge. For instance, Cannon-Bowers and Salas (1997) have recently proposed a framework that integrates many aspects of team cognition in the form of teamwork competencies. They categorize competencies required for effective teamwork in terms of knowledge, skills, and attitudes that are either specific or generic to the task and specific or generic to the team. Similarly, a team's understanding of a complex and dynamic situation at any one point in time (i.e., team situation awareness) is supposedly influenced by the knowledge that the team possesses (Cooke, Stout, & Salas, 1997; Stout, et al., 1996). These distinctions between long-term knowledge and knowledge that is associated with a dynamic or fleeting understanding of the situation are reflected in our framework, as are distinctions between knowledge about the task (i.e., taskwork knowledge) and knowledge about the team (e.g., teamwork knowledge). We also acknowledge that there are many other distinctions that could be made in regard to knowledge type (e.g., declarative, procedural, strategic) which are not reflected in this framework.

Another critical distinction that is depicted in the framework of Figure 1 is among different levels or metrics of team knowledge. Individual team members have knowledge that can be aggregated or summed. We use the term “collective” to portray the aggregation of individual knowledge. Collective metrics are distinct from holistic metrics of team knowledge. In our view, the team knowledge resulting from team process behaviors is “effective” knowledge that guides team action and that is, therefore reflected in outcome measures. Eliciting or assessing team knowledge at this *holistic* level is analogous to focusing on an individual's effective knowledge revealed in his or her actions. Thus, team knowledge at this holistic level is team member knowledge that has been processed or integrated in some way through team behaviors such as communication, coordination, or leadership. For example, team members in a military aviation setting may *collectively* have information about an impending threat, but without adequate communication, and therefore integration or fusion of the requisite pieces of information, the *holistic* knowledge would be lacking and the team would fail to act on the impending threat. Holistic metrics of team knowledge more directly reflect our definition of team cognition in that it emerges from the interplay of individual team knowledge and team process behaviors.

Although specific characteristics of our framework may be novel, the proposal that team cognition exists is not. The concept of transactive memory assumes that two or more individuals can share the task of remembering by using each other as memory storage components (Wegner, 1986). Further, concepts like distributed cognition (Hutchins, 1991) and common ground in discourse (Clark & Schaefer, 1987; Wilkes-Gibbs & Clark 1992) also presume that cognitive tasks can be shared among two or more individuals. There has also been significant theoretical work delineating cognitive constructs at the team level such as team decision making, shared mental models, and team situation awareness (Cannon-Bowers, Salas, & Converse, 1993; Orasanu, 1990; Stout, et al., 1996). The interest in such concepts is based on the hypothesized relation between team cognition and team performance that suggests that team performance can be better understood and ultimately predicted by measuring and understanding team cognition. We agree that these team cognition constructs are key to understanding team performance. However, when we examine the way in which these constructs are measured in light of our framework (see Figure 1) we identify a number of measurement limitations. These limitations and our approaches to address them are the topic of the remainder of this chapter.

Limitations of Measures of Team Cognition

The assessment and understanding of team cognition (i.e., team mental models, team situation awareness, team decision making) requires psychometrically sound measures of the constructs that comprise team cognition. However, measures and methods targeting team cognition are sparse and fail to address some of the more interesting aspects of team cognition (Cooke, Salas, Cannon-Bowers, & Stout, 2000). In particular, we have identified several issues that need to be addressed in order to adequately measure team cognition. These measurement needs are presented in Table 1 and will be discussed in turn.

[Insert Table 1 about here]

Holistic vs. Collective Measurement

Traditional measures of team knowledge (e.g., measures of shared mental models) are based on the elicitation of knowledge from individuals on the team, which is then aggregated (e.g., averaged) to generate a representation of the team's knowledge. Thus, this aggregate represents the "collective" level of Figure 1and thus, is devoid of the influences of team process behaviors (e.g., communication, coordination, situation awareness).

To what extent does team knowledge differ when measured at the collective versus the holistic level? How do collective and holistic knowledge compare in terms of their ability to predict team performance? Based on the framework in Figure 1, we predict that holistic knowledge, by virtue of its direct link to team performance, would be a better predictor of team performance compared to collective knowledge. Can team knowledge measured at the collective level be transformed in some way to better reflect the holistic level? How can knowledge be elicited or measured at the holistic level? We have begun to address this latter question that we view as a prerequisite to answering the others.

Heterogeneous vs. Homogeneous Measurement

Traditional measures of team cognition or team knowledge reflect an underlying assumption of homogeneous cognition or knowledge among members of a team. For instance, many measures of shared mental models are based on averaging responses elicited from individual team members. Not only is aggregation focused at the collective, rather than the holistic level, but typical aggregation metrics such as averaging assume that what is being measured (e.g., taskwork knowledge) varies only quantitatively. That is, knowledge is a matter of degree or the extent to which a core set of facts is mastered. Further, typical measures of shared mental models are based on an assumption that shared knowledge among team members is the same as similar knowledge among team members.

In contrast, it is possible to conceive of team members as a collection of "apples and oranges," each having very different knowledge, skills, and attitudes that they bring to the task. Indeed, this characterization of team members is central to a common definition of team offered by Salas, Dickinson, Converse, and Tannenbaum (1992).

They define *team* as "a distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common and valued goal/object/mission, who have each been assigned specific roles or functions to perform, and who have a limited life span of membership" (p. 4). Thus, teams, unlike some groups, have differentiated responsibilities and roles (Cannon-Bowers, Salas, & Converse, 1993). This division of cognitive labor is quite common (e.g., surgery, battlefield command) and enables teams to tackle tasks too complex for any individual. In this context, shared knowledge may be likened to sharing a piece of pie (i.e., each team member has a different, but complementary piece), rather than sharing in the sense of a belief or value (i.e., in which the belief or value is identical for all team members). This heterogeneous nature of teams is a feature that has been neglected by current measurement practices (e.g. Langan-Fox, Code, & Langfield-Smith, 2000). It also poses a challenge to measuring team knowledge or shared mental models.

Need for Improving Aggregation Methods

Traditional aggregation procedures (e.g., averaging) associated with collective metrics assume homogeneity with respect to team member knowledge. For instance, a good example of a team for whom a simple averaging of individual scores is inappropriate is a structured team, in which every member has a different role. In an operating room we would not simply average scalpel-wielding performance, as the anesthesiologist would likely have a 0, regardless of how well she anesthetizes. There is a need for more principled aggregation approaches suited to heterogeneous teams such as process-oriented methods for aggregating individual knowledge to attempt to predict holistic knowledge (e.g., the social decision scheme literature, Hinsz, 1999; Davis, 1973).

Essentially, these strategies consist of assessing how a team might combine their individual scores into a team score, based on what we know about their team process.

Dynamic Fleeting Knowledge vs. Long-term static Knowledge

Also, as portrayed in Figure 1, we conceive of team knowledge as both background knowledge that is long-lived in nature, and a more dynamic and fleeting understanding that a team member has of a situation at any one point in time. Indeed, many applied researchers have highlighted the importance of this latter kind of cognition, often termed situation awareness (Durso & Gronlund, 1999; Endsley, 1995a; Sarter & Woods, 1991). There has also been a growing appreciation of the importance of team situation awareness or that rapidly changing information that a team uses to understand and make decisions in a rapidly evolving environment (Cooke, Stout, & Salas, 2001; Prince, Stout, & Salas, 1994; Salas, Prince, Baker, & Shrestha, 1995). This dynamic, fleeting knowledge is a challenge to measure at the individual level compared to the more stable knowledge associated with most knowledge elicitation methods (Adams, Tenney, & Pew, 1995; Endsley, 1995b, Fracker, 1991, Salas, et al., 1995; Sarter & Woods, 1995; Taylor & Selcon, 1994; Vidulich, 1989). Thus, measures of team knowledge have focused primarily on long-term team knowledge, at the expense of the more dynamic knowledge associated with situation awareness.

Need for Broader Focus on Different Knowledge Types

The measurement of team cognition has tended to focus on the knowledge constructs of shared mental models (Cannon-Bowers, Salas, & Converse 1993; Kraiger & Wenzel 1997; Klimoski & Mohammed, 1994) and to a lesser extent team processes involved in team decision making and communication. As previously mentioned, there have been a few recent attempts to measure knowledge associated with situation awareness. However, even within the arena of long-term team knowledge (and situation awareness as well) there are a number of other distinctions that have largely been ignored. For instance, Cannon-Bowers , Tannenbaum, Salas, and Volpe (1995) distinguish between taskwork and teamwork knowledge. Taskwork knowledge is knowledge about the individual and team task and teamwork knowledge is knowledge about the roles, requirements, and responsibilities of team members. Others have distinguished between strategic, procedural, and declarative knowledge (Stout, et al., 1996). These theoretical distinctions have yet to be captured by measures of team knowledge, which should in turn, put these distinctions to the test.

Need for Broader Application of Knowledge Elicitation Methods

Cooke (1994) outlines the various knowledge elicitation techniques that have been applied to individual knowledge elicitation and Cooke, et al. (2000) describe how these techniques can be applied to knowledge elicitation at the team level. In summary, the varieties of knowledge elicitation methods can be categorized into one of four categories: 1) observations, 2) interviews and surveys, 3) process tracing, and 4) conceptual methods. Those methods that have been used to elicit team knowledge relevant to shared mental models or situation awareness have been typically restricted to interview and survey methods and to a lesser extent, conceptual methods. The application of a broader spectrum of knowledge elicitation methods to the problem of measuring team knowledge could open the door to the measurement of varieties of team knowledge previously untapped. As will be described in what follows, we have begun to explore the application of process tracing techniques at the team level through the opportunistic use of team communications.

Need For Automated and Embedded Measures

There are several reasons why it is particularly advantageous to have automated and embedded measures in team research. Most of them trace back to the fact that applied research on team tasks tends to be more complex and have more variables and measures. As task complexity increases, it becomes more disruptive to interrupt and administer a questionnaire or interview an expert. Moreover, as the size of the team increases, the time needed to query all team members increases. So, essentially, even fairly efficient measures can become difficult to administer and retain when the flow of the task is more fragile, when there are multiple team members being queried, and when multiple measures are being taken. Automating measures and embedding them within the task can reduce task disruption due to measurement, along with experimenter measurement errors and experimenter resources.

Need for Validation

The units used in team research are typically not items on a questionnaire, but rather single team-level measures. It is therefore unlikely that there will be enough observations to conduct certain standard psychometric tests, such as Cronbach alpha for reliability. On the other hand, test-retest reliability is a possibility if there are multiple trials on which the measure is taken. Also, if the measure is the judgment of more than one person, then interrater reliability remains viable.

Two approaches are readily taken to assess validity. First, if there are repeated trials in which a measure is taken, it can be determined if the measure shows an acquisition curve that corresponds to the team's learning of the task. The measure should show either a power curve or an inverse linear curve. The second strategy is to show criterion validity to some other measure, such as performance.

Some New Approaches to Measuring Team Cognition

In this section we describe the approaches that we have taken in an effort to begin to address the limitations of measures of team knowledge described above.

Holistic Knowledge Measurement

The approach we have taken to holistically assessing knowledge is to have the team collaborate on their responses to queries that would otherwise have been administered individually. For instance, when we ask individuals to make pairwise relatedness ratings of pertinent concepts in the task, we immediately follow this procedure with asking the team to make their ratings, this time reaching a verbal consensus on each value they assign. For each pair, the rating entered in the prior session by each team member is displayed on the computer screen of that team member. The three team members discuss each pair until consensus is reached.

We have assumed that this assessment at the team level would capture not only the collective knowledge of the team members, but also process behaviors of the team that are used in coming to consensus on the ratings. So, for instance, if a team has a particularly persuasive-- but wrong-- team member, then their team score is accurately reflected as being too close to that misinformed score. The same process takes place during actual task performance, for better or for worse. Because of the inclusion of team process in the ratings, it was hypothesized that the consensus ratings would be better predictors of team performance than the aggregate taskwork ratings.

On the other hand, it is possible that the consensus measure is not a good estimate of holistic knowledge. The conformity demanded by the measure may induce process behavior different from task-related process and the public nature of the consensus may overshadow private disagreements. Whether such discrepancies result in output that is a good predictor, or at least a better predictor of performance than collective measures, awaits empirical results.

We have devised similar holistic rating procedures for measures of fleeting situation awareness knowledge, for knowledge of interaction among team members (i.e., teamwork knowledge), and for knowledge of the functioning of the task itself (i.e., taskwork knowledge). Order of presentation between holistic and individual cannot be randomized, due to the carryover effects of achieving consensus. So the challenge is to keep the task short enough to minimize fatigue or boredom effects.

Metrics for Heterogeneous Teams

To assess knowledge of heterogeneous teams we have computed knowledge accuracy metrics that are role-specific. A referent or "answer key" is generated by the experimenters for each role, by only considering what knowledge a team member in that role should be expected to know, in order to do their job well. Individual query responses are scored not only against a key representing overall knowledge, but also against role-specific keys. In this way, measures of "role" or "positional" accuracy, as well as "interpositional" accuracy (i.e., interpositional knowledge (IPK) or knowledge of roles other than their own) can be determined. Effective performance can ultimately be characterized in terms of degree of knowledge homogeneity among team members.

Aggregation of Team Data

If measures are not collected at the team or holistic level, but at the individual level, then how can the individual measures best be used to estimate holistic, team-level scores? Here we label any *a priori* aggregation scheme, whether based on central tendency (i.e. attempting to choose the "typical" score for the team), dispersion (i.e. assessing dissimilarity among team members), or anything else, "collective." A more empirically-driven collective approach attempts to find a function that combines individual scores to best predict team scores. We are in a particularly good position to do this when we have data collected both at the individual and at the holistic (team consensus) level.

Our framework for thinking about team cognition (see Figure 1), as well as work of others (e.g., Steiner, 1972) suggest that we should attempt to correct our collective estimates so that they reflect change in individual scores due to process loss (or possibly gain). For instance, one could compute central tendency corrected for dispersion by making aggregation contingent upon the result of cluster analysis (Everitt, 1993), or other similarity methods. Aggregation methods based on team process are relatively ideographic, in that there will be a different combination rule for every team, depending on what their process is like.

Social decision scheme theory (SDS, Davis, 1973; Hinsz 1999; Hinsz, in press) takes this ideographic attempt to estimate holistic scores as its explicit focus. SDS attempts to combine individual decision making in such a way that it predicts the actual team decision among discrete alternatives. Combination rules are proposed based on theories about how specific groups interact, rather than on a single nomothetic rule applied to all teams, as the traditional collective approach does. Theories about team process are tested by determining how well the theoretical combination of scores under an SDS matches the actual decision.

In our work, in order to identify strategies that teams use to come to consensus in a knowledge rating task, the three individual and one team response for each item within a knowledge measure is examined as a single response set. Each response set can be classified according to one of five rules that map individual responses onto the team response. For example in the context of the pairwise relatedness ratings used to measure taskwork knowledge we can identify response sets in the following categories (TM = team member:

1) all agreed (e.g., TM1=5, TM2 = 5, TM = 5, Team = 5)
2) majority (2 out of 3) rules (e.g., TM1 = 4, TM2=4, TM3 = 3, Team =4)
3) leader emerges (e.g., TM1=3, TM2=0, TM3=1, Team =3 or TM1=4, TM2=4, TM3 = 2, Team =2)
4) mid rating (e.g., TM1=0, TM2=3, TM3=5, Team =2 or TM1=0, TM2=3, TM3=5, Team=3)
5) different from each, and not middle rating (e.g., TM1=5, TM2=2, TM3=4, Team=0)

Process strategies identified through these or similar data could then be used to aggregate other individual measures taken on the same teams. Interestingly, in our analyses we found that teams used a variety of strategies, however, there was no connection between the strategy they used and team performance. In particular teams used *majority rules* and the *leader emerges* strategies frequently. In addition, they used the fifth “nonstrategy” fairly often, which led us to question the degree to which team members took the consensus rating task seriously.

Measuring Broader Varieties of Team Knowledge

In our approaches to measuring team knowledge we discriminate among varieties of knowledge that we target. For instance, we have developed some strategies for measuring team situation awareness, or more specifically, the dynamic situation model that teams create. This is distinguished from the longer-term background knowledge that experienced team members bring to the task. Within the long-term knowledge category, we have additionally distinguished methods that elicit teamwork knowledge from those that elicit taskwork knowledge. In this section we describe these efforts.

We, like others (e.g., Endsley, 1995a) view the team process of situation assessment as distinct from the outcome of this process, which is the knowledge that is embodied in the team situation model. It is the latter that we focus on here. We view the team situation model as the team’s understanding of a specific situation at any one point in time (Cooke, Stout, et al., 2001). As suggested in Figure 1, the situation model is a product of situation assessment and other team process behaviors and individual situation models of each team member.

In our measurement of the team situation model we have extended situation awareness query methods such as SAGAT (Situation Awareness Global Assessment Technique; Endsley, 1990) and SPAM (Situation Present Assessment Method; Durso, Hackworth, Truitt, Crutchfield, Nikolic, & Manning, 1998) to the team situation. Most of our measures have relied on the SPAM method in which the environmental cues remain present (i.e., the display is not frozen or blank as required by the SAGAT method) and queries focus on future events or present events not immediately apparent from perusal of a single display.

We typically query each individual on the team about aspects of the task environment in the present or future (e.g., how many targets will your team manage to successfully photograph by the end of this mission?). Queries are either triggered by events or presented at randomly determined times. Individual responses are either written or oral and in the case of oral responses, team members are queried in random order within a five-minute interval. Team accuracy can be estimated through aggregation of individual accuracy scores, though recently, we have begun to present queries to the team as a whole for a consensual response. We also examine the similarity or agreement of individual responses as another measure of the team situation model. In addition to implementing a holistic procedure, we have recently addressed the issue of repeating the same situation awareness queries across multiple missions.

We have explored a number of off-line (apart from the task) methods of assessing taskwork and teamwork knowledge of individuals and the team as a whole. Although we have assessed taskwork knowledge using multiple choice test items pertinent to the task and by having team members analyze the task in terms of goals, subgoals, and objectives, the most successful approach thus far has relied on pairwise relatedness estimates of task-relevant concepts. That is, individual team members rate concepts like *airspeed* and *altitude* in terms of their overall relatedness in the task context. The concepts, which typically number 10 to15, are carefully selected as representative of critical task concepts. In addition, selected concept pairs are hypothesized to discriminate between team members with low and high taskwork knowledge or between team members with different task roles. After rating data are collected, the results can also be used to further identify pairs that empirically discriminate individuals and teams.

We have used Pathfinder network scaling as a means of meaningfully reducing the quantitative ratings to a graphical structure (Schvaneveldt, 1990). The Pathfinder procedure also offers a means of comparing two or more resulting network structures quantitatively, thereby allowing comparisons of taskwork knowledge of two or more individuals or teams (aggregate networks based on average ratings of team members). This also allows comparisons of team or individual networks to a referent to determine knowledge accuracy. As mentioned previously, we have also extended this particular measure of taskwork knowledge to allow for holistic measurement and to address heterogeneous teams.

Our approach to measuring teamwork knowledge has also focused (until recently) on individual measurement and aggregation. In general, we elicit from individual team members their understanding of information requirements within a given scenario. The measure has evolved from a task in which team members used index cards to represent team members and information links between team members to a questionnaire in which team members read a scenario and check relevant items in a list where each items refers to a team member passing a specific type of information to another team member. Most of our data on the teamwork measure, however, come from a more open-ended questionnaire in which team members first indicate who is talking to whom in a given scenario and then describe the content of the communication. This measure can also be scored for intrateam similarity as well as accuracy.

Using Communication as a Window to Team Cognition

We have proposed a general methodological approach for semi-automatically assessing team cognition using communication data (Kiekel, Cooke, Foltz, & Shope, 2001). The approach rests on the premise that analyzing communication data is a means of assessing team cognition. If teams are to be the unit of analysis under our holistic definition, we will need to measure behaviors exhibited by the team as a whole. Just as we use think-aloud verbal protocols to make inferences about individual cognition, so we have used the communication inherent in the collaboration process to assess team cognition in a holistic way. In this respect, team cognition is easier to measure than individual cognition; teams need not interrupt their process in order to think-aloud, since they are always "thinking aloud" in some sense. With a newly formed team, team cognition begins as the sum of individual cognition. Then, as the team thinks (interacts), dynamic changes occur in the team mind as a natural result of the interaction. Effects of this process on performance depend on the type of task (Steiner, 1972). Thus, an analysis of team communication provides a window through which to view team cognition.

The need for automation in measurement and analysis is sparked by the time-consuming process of achieving reliable results with communication data. Such data are often voluminous and/or rich in contextual information. We have approached two aspects of communication data, communication flow from team member to team member and communication content. For flow, we have recorded frequency of speech by each team member at each second, or at smaller specified intervals. We have developed and utilized several methods-- especially sequential methods-- for addressing these data. For instance, we have developed CHUMS (Clustering Hypothesized Underlying Models in Sequence; Kiekel, Cooke, Foltz, Gorman, & Martin, 2002) to cluster adjacent time units, so we can look for pattern changes (e.g. in amount of speech by each person). Another example is our use of ProNet (Procedural Networks; Cooke, Neville & Rowe, 1996) to determine which speech events tend to co-occur in time (e.g. Pilot begins speaking--Navigator interrupts--Pilot finishes speaking).

To analyze content data, we have quantified the actual content of the speech, both manually and using LSA (i.e., Latent Semantic Analysis; Landauer, Foltz, & Laham, 1998). LSA is a model of human language that can be used to code content and to determine similarity among utterances. It works by assessing co-occurrence of words within a large representative corpus of the language (e.g. encyclopedias, text books, transcripts). It first defines a matrix of the frequency of co-occurrence of each word with each other word. Co-occurrence is defined as two words occurring in the same paragraph. Then this rudimentary knowledge of the language is generalized to semantic knowledge. LSA extracts latent dimensions underlying the co-occurrence matrix by SVD (Singular Value Decomposition, the machinery behind principal components analysis and factor analysis). LSA then retains the first 300-400 dimensions (however many the user chooses), so that each word is quantified by a vector in 300-400 dimensional space. Each utterance is a vector made of the mean of its words. These vectors each have length, which is akin to the amount of meaning in an utterance. Furthermore, utterances can be correlated by taking their cosine. Using LSA, we have looked at variables such as the continuity of similarity from one utterance to the next, similarity of the whole team's dialogue to other teams whose performance score was known, and overall internal consistency of the team's communication, defined by the mean of the similarity matrix among all utterances.

Automating and Embedding Measures

We have approached the goal of automated and task-embedded knowledge measures along several fronts. There is still much to do. For instance, our measures of taskwork and teamwork knowledge are still collected off-line. The teamwork measures are paper-based questionnaires, and although the taskwork relatedness ratings are collected via computer, the sheer number of ratings required for a pairwise ratings task where N is equal to the number of concepts is (N x (N-1)) /2. This translates to 45 ratings for 10 concepts and 190 for 20. Consequently, as concepts are added, the task becomes tedious quickly and requires streamlining.

Alternatively, our team situation awareness measures represent progress along this front. We are firm believers in embedded measures of situation awareness, as it is necessary to measure that knowledge in the context of the situation at the moment when momentary understanding occurs. So, we have successfully embedded our team situation awareness measures in the task and have even generated somewhat automated and on-line probe sheets as cues for the experimenters.

Possibly our greatest contribution in the area of automation and task-embedded measures has resulted from the work described previously in which we use team communications as an automatically-generated and task-embedded think aloud task at the team level. We can then infer team cognition or team knowledge from this team think aloud.

Validating New Measures of Team Cognition

Our approach has been to validate the measures described previously through experimentation in which we examine the changes in the measure with skill acquisition and the degree to which the measure is predictive of performance. Experimentation occurs in the context of STEs (Synthetic Task Environments), laboratory abstractions of operational tasks such as Navy helicopter rescue missions (Cooke, 1998; Cooke, Kiekel, & Rivera, 2000; Cooke, Kiekel, Salas, Stout, Bowers, & Cannon-Bowers, in press) and UAV (Uninhabited Air Vehicle) ground operations (Cooke, Kiekel, & Helm, 2001; Cooke, Shope, & Kiekel, 2001).

Looking at the changes in acquisition curves across trials provides at least two pieces of information. First, it is a means of validating the measure, as measures of knowledge or skill gained in the course of the experimental trials should show a learning curve with repeated trials and covary with skill acquisition. Second, testing post-asymptote trials against one another can assess reliability for any given measure. If it is known that the team is at asymptote, then any variation in the same team on the same score must be due to unreliability.

Second, team process and cognitive or knowledge measures can be judged for validity based on how well they predict performance. Whether the prediction needs to be a direct or an indirect prediction of performance-- i.e. a path that leads through another variable-- depends on the constructs under consideration. Further when team performance is affected by manipulations of various types (e.g., workload changes, communication mode changes), it is expected that such variations in performance should also be reflected in underlying cognitive and behavioral constructs that support this performance. For example, when workload demands of the task are increased, team performance typically declines, and this decline should similarly be apparent from the measures of relevant cognitive and behavioral constructs.

What Have We Learned About Team Cognition?

We have conducted several team experiments in which we measure team knowledge in addition to team performance and process (Cooke, 1998; Cooke, Kiekel, et al., 2001; Cooke, et al., 2000; Cooke, et al., in press; Cooke, Shope, et al., 2001). We measure team performance using composite scores based on task outcomes and we measure team process behaviors using observer ratings of process behaviors along dimensions such as communication, coordination, and decision making or by using a behavioral checklist in which target behaviors are associated with event-triggers in the context of the task. These performance and process measures serve as criterion values for our team knowledge predictors. Newly developed or modified measures of team knowledge are then applied in these experiments and evaluated in terms of their ability to predict team performance and process and to reflect any other experimental manipulations thought to influence team knowledge such as team task experience and training regime.

As previously mentioned, our experiments were performed in the context of synthetic task environments. Findings that were replicated in at least three of the four studies are reported below. Note that although these findings appear to generalize across our two synthetic tasks, the extent to which they generalize to other teams, team tasks, and work domains is an empirical question for further research.

Taskwork Knowledge is Predictive of Team Performance

Across our studies we have generally found that taskwork knowledge, measured after individual task training, is predictive of team performance. That is, teams with members who understand the task immediately or soon after training tend to be high-performing teams. For example, in one of our UAV studies (Cooke, Kiekel, et al., 2001), taskwork knowledge measured after training and the first mission predicted team performance on the second mission ($r(9) = .839$, $p<.01$) and the tenth mission ($r(8) = .725$, $p<.05$). However, there are a few caveats to this generalization, even within the context of our studies.

First, this finding is true when taskwork knowledge is measured using taskwork relatedness ratings (i.e., team members rate pairs of task relevant concepts for relatedness and these ratings are analyzed using the Pathfinder network scaling algorithm) and to a lesser extent, the consensus version of the ratings targeting holistic taskwork knowledge. For instance, in the same UAV study mentioned previously (Cooke, Kiekel, et al., 2001), taskwork knowledge accuracy as based on the individual taskwork ratings predicted team performance on the first four missions with correlations ranging from .535 to .839 (df = 9), whereas when taskwork knowledge accuracy is based on the consensus ratings correlations with team performance on the same missions ranged from .352 to .549 (df = 9). Accuracy using these rating measures is based on comparison to a referent network indicating pairs that should or should not be related. The ability of taskwork knowledge to predict team performance, however, does not hold when measured more directly using multiple choice tests or unstructured questionnaires requiring goal and task decomposition.

This does not imply that these more direct methods are invalid. Alternatively, it may be that our implementation of them is flawed. Indeed, our multiple choice questions have tended to be too easy and thus, not very sensitive, whereas the goal and task decomposition exercise was so open-ended that it produced a wide range of mainly terse and vague responses that were very difficult to score. So, although we have had the most success with the relatedness rating task, team knowledge of the task may be measured adequately using other methods implemented in a different way. On the other hand, we believe that the relatedness ratings task works (i.e., is sensitive to team performance differences) because of its indirect nature. Judgments are made more easily than explicit facts are recalled, and seemingly minor discrepancies in some judgments may reflect the difference between an expert and a novice (Schvaneveldt, Durso, Goldsmith, Breen, Cooke, Tucker, & DeMaio, 1985). Also, in many ways, the similarity of two sets of relatedness estimates, or two resulting networks, seems more in line with the idea of shared mental models. Although we would *not* expect two individuals with "shared mental models" to know the identical facts and rules in regard to a task (as measured by more direct tests of knowledge), it seems reasonable that they would have the same perspective on a task, reflected in similar relatedness judgments.

The second caveat to the finding that taskwork knowledge is predictive of team performance is that with the application of our heterogeneous knowledge metrics we are getting a better idea of the precise nature of that knowledge. In the UAV context, teams that are high-performing have members who each understand the task as a whole (i.e., the big picture) and from the perspective of each team role or position. For instance, for good UAV teams, the Air Vehicle Operator understands the task from all perspectives in that he or she provides ratings that are accurate in regard to the Air Vehicle Operator referent, as well as the referents associated with the Payload Operator, and Navigator (i.e., DEMPC) roles. In a UAV study for instance (Cooke, Kiekel, et al., 2001), correlations between team performance and interpositional taskwork knowledge based on ratings and comparisons to role referents ranged from .232 to .613 for the first four missions and .555 to .677 for the last three missions. In other words, good teams have interpositional knowledge of taskwork. So, for these highly interdependent team tasks that we have studied, it is not just a matter of becoming expert at a single task, but also of understanding the other roles well enough to provide “accurate” relatedness ratings of task concepts relevant to these other roles.

Finally, one other finding has emerged that has direct support in one of the helicopter studies and indirect support in several other studies. Specifically, there appears to be a sequential constraint on the acquisition of taskwork versus teamwork knowledge. In a cross training study (Cooke, et al., in press), we attempted to short-cut traditional cross training in which all team members are given complete training in each role. We compared this full cross training condition which emphasized taskwork to a condition in which team members were cross trained not in the other tasks themselves, but on teamwork information. That is, information about the responsibilities and information requirements associated with each team role was conveyed through notes and charts. We hypothesized that because this latter condition preserved what we believed to be the essence of cross training, it should take less time and be at least as successful in terms of team performance relative to full cross training.

Contrary to our expectations, full cross-trained teams showed marginal benefits in terms of team performance, but significant taskwork and teamwork knowledge benefits. That is, full cross-trained teams had more interpositional knowledge than other teams in regard to both taskwork and teamwork knowledge (see Cooke, et al., in press for details). Further analyses suggested that the acquisition of teamwork knowledge (trained explicitly in the abbreviated training condition) was facilitated by a prior understanding of taskwork knowledge (conveyed in the full cross training condition). In other words, team members need to understand what is done before they understand who does it. But, why did we see only marginal performance benefits with full cross training? It appears that for this task, unlike the UAV task, taskwork knowledge specialization, not interpositional knowledge, is associated with high performing teams. Thus, cross training may be at odds with the knowledge requirements of this task. These interesting possibilities regarding knowledge requirements and sequential constraints in the acquisition of taskwork and teamwork knowledge were only revealed through our measurement of team knowledge, and thus provide support for the diagnostic power of the measurement of team cognition.

Team Situation Awareness is Predictive of Team Performance

In our studies we have also found that team SA (situation awareness) is a consistently good predictor of team performance. Teams with high SA tend to be better performing teams. For instance in a UAV study (Cooke, Kiekel, et al., 2001), team accuracy scores on situation awareness probes (averaged across 10 missions) were correlated with team performance (also averaged across the 10) missions; $r(11) = .88$, $p < .0001$). In addition, team situation awareness is one type of knowledge that continues to improve after training with team task experience. For instance, Cooke, Kiekel, et al. (2001) found that team situation awareness accuracy and intrateam similarity improve across missions and parallel skill acquisition on the task. This might be expected as the knowledge and expectations that are tied to specific mission situations cannot be conveyed in our factual tutorials as well as through direct exposure to the situation.

In some cases the correlation between team SA and team performance have been so high (e.g., the *r* (11) = .88 referenced above) that we have questioned the validity of our measure. It is possible that through repeated exposure to SA queries, team members acquire situation awareness that focuses on information relevant to the queries, thus artificially inflating their SA accuracy. However, if this is the case, then it is the acquisition of this "test knowledge," rather than SA that is associated with high team performance. Indeed, our most recent results suggest that accuracy on repeated SA queries is more highly correlated with team performance that accuracy on non-repeated queries. This suggests that the SA that we measure that is developing and mirroring performance is constrained to the specific task and experimental situation.

Teamwork Knowledge Changes with Experience

In our studies we have also seen a general tendency for the teams' knowledge of teamwork to increase over time. For example, Cooke, Kiekel, et al. (2001) found that accuracy on the teamwork questionnaire increased over the four knowledge sessions (.53, .66, .71, and .65) with the largest increase paralleling the team performance acquisition function. On the other hand, discrepancies in the accuracy of teamwork knowledge do not predict team performance and thus, this is something that teams seem to acquire uniformly with team task experience. The general pattern seems to be that the team's knowledge of taskwork is acquired early through the training tutorials (largely declarative knowledge), but this is not acquired to the same degree by all teams and discrepancies are predictive of team performance. The more fleeting knowledge (i.e., SA) and the team's knowledge of teamwork are acquired through actual task experience (largely procedural knowledge). Also, discrepancies in SA, but not teamwork knowledge predict performance. The sequential constraint discussed previously in regard to teamwork and taskwork knowledge may explain the different patterns of knowledge acquisition. That is, it seems important for the acquisition of taskwork knowledge to precede teamwork knowledge and so unlike SA, it may not be that exposure to actual mission situations is critical to learn about teamwork. Rather the sequential constraint demands that the acquisition of teamwork knowledge wait until taskwork knowledge is mastered. In addition, it is possible that our teamwork measure (questions about information flow in the team) is not sensitive enough to reveal subtle differences among teams that are reflected by the SA and taskwork measures.

Communication Consistency Predicts Team Performance

We believe that the use of communication data as a window to team cognition is not only theoretically justified, but also practical given that these data can be collected unobtrusively during mission performance. The challenge is in finding methods of communication analysis that best reflect team cognition. Thus far we have had promising results using LSA (Landauer, et al., 1998) to identify similarity in segments of discourse based on semantic content. We have also used ProNet (Cooke, et al., 1996) to examine patterns of communication flow among team members across a 40-minute mission. In both cases, the preliminary results (Kiekel, et al., 2001) suggest that teams who perform poorly exhibit a variety of communication patterns both in terms of flow and content. This variation occurs within individual teams. On the other hand, high-performing teams demonstrate consistency over time in both what they say (i.e., content) and who they say it to and when (i.e. flow). In general, the results of the semi-automated approaches to communication analysis have been successful in terms of predicting team performance. Next steps involve associating communication patterns with team characteristics or behaviors more diagnostic of cognitive or behavioral deficits or strengths (Kiekel, et al., 2002).

Concluding Remarks

In this chapter we have presented our motivation for investigating team cognition, have discussed its importance, and have offered a definition of team cognition. In light of this definition, we have highlighted several limitations of traditional measurements of team cognition. Our recent research in the area begins to address some of these limitations by developing new methods or adapting older methods for measuring team cognition. Of primary importance are the development of automated and task-embedded measures and the evaluation of new measures in terms of their validity. Thus far, we have identified some valid measures of team cognition, specifically taskwork knowledge, teamwork knowledge, and the knowledge associated with team situation awareness. In four studies we find that taskwork knowledge seems to precede the acquisition of teamwork knowledge and that teams with members who understand the task from the perspectives of roles other than their own (assessed using metrics for heterogeneous teams) tend to be high-performing teams. We have also begun to develop and test holistic measures of team knowledge that approach measurement by eliciting or assessing knowledge at the team level. Finally, we are investigating team communication analysis methods because they offer the promise of automated and task-embedded measures of team cognition.

This chapter documents some small steps taken in an effort to help understand fully what comprises effective team functioning. Team cognition represents the most challenging and difficult pillar for providing a comprehensive picture of the attitudes, behavior, and cognition of teams. It is only, we believe, through theoretically-driven systematic lab and field studies that focus on designing, developing and validating sensitive, diagnostic and useful measures of team cognition that we will contribute a much needed answer to the numerous businesses, industries and agencies for which teams are central. We hope this chapter and the others in this volume motivate more thinking and research on this important topic.

Acknowledgements

This work was supported by AFOSR grant F49620-01-0261 and ONR grant N00014-00-1-0818 to the lead author and by AFOSR grant F4620-01-0214 to the second author. The authors would like to thank Janie DeJoode, Jamie Gorman, Heather Jiron, Rebecca Keith, Harry Pedersen, Steven Shope, and Natasha Sutherlin, for their contributions to the research described here. Thanks also to the helpful suggestions made by an anonymous reviewer.

References

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors, 37,* 85-104.

Cannon-Bowers, J. A., and Salas, E. (1997). Teamwork competencies: The interaction of team member knowledge skills and attitudes. In O. F. O'Neil (Ed.), *Workforce readiness: Competencies and assessment* (pp. 151-174). Hillsdale, NJ: Erlbaum.

Cannon-Bowers, J. A., & Salas, E. (Eds.). (1998). *Making decisions under stress: Implications for individual and team training.* Washington, DC: APA.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. Castellan Jr. (Ed.), Current *issues in individual and group decision making* (pp. 221-246). Hillsdale, NJ: Erlbaum.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Teams: Their training and performance* (pp. 101-124). Norwood, NJ: Ablex.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81.

Clark, H. H., & Schaefer, E. F. (1987). *Collaborating on Contributions to Conversations. Language and Cognitive Processes, 2,* 19-41.

Cooke, N. J. (1998). Measures of Team Knowledge. Report for Purchase Order No. N61339-97-M-0507, submitted to Naval Air Warfare Center Training Systems Division, Orlando, FL.

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41,* 801-849.

Cooke, N. J., Kiekel, P. A., & Helm E. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics: Special Section on Knowledge Acquisition, 5,* 297-315.

Cooke, N. J., Kiekel, P. A., & Rivera, K. (2000). Conceptual Cross Training for Teams: Improving Interpositional Knowledge. Technical Report submitted to U.S. Army Research Office, TCN 98020, Contract DAAH04-96-C-0086.

Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C, & Cannon-Bowers, J. (in press). Measuring team knowledge: A window to the cognitive underpinnings of team performance. Submitted to *Group Dynamics*: *Theory, Research, and Practice.*

Cooke, N. J., Neville, K. J., & Rowe, A. L. (1996) Procedural network representations of sequential data. *Human-Computer Interaction, 11,* 29-68.

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

Cooke, N. J., Shope, S. M., & Kiekel, P.A. (2001). Shared-Knowledge and Team Performance: A Cognitive Engineering Approach to Measurement. Technical Report for AFOSR Grant No. F49620-98-1-0287.

Cooke, N. J., Stout, R., & Salas, E. (1997) Expanding the measurement of situation awareness through cognitive engineering methods, *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting,* (pp. 215-219). Santa Monica, CA: Human Factors and Ergonomics Society.

Cooke, N. J., Stout, R.J., & Salas, E. (2001). A knowledge elicitation approach to the measurement of team situation awareness. In M. McNeese, M. Endsley, & E. Salas, (Eds.), *New Trends in Cooperative Activities: System Dynamics in Complex Settings,* pp. 114-139. Santa Monica, CA: Human Factors.

Davis, J. H. (1973). Group decision and social interaction: A theory of social decision schemes. Psychological Review, 80(2): 97-125.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J., & Nikolic, D. & Manning, C. A. (1998). Situation awareness as a predictor of performance in en route air traffic controllers. *Air Traffic Control Quarterly, 5,* 1-20.

Endsley, M. R. (1990). A methodology for the objective measure of situation awareness. In *Situational awareness in aerospace operations* (AGARD-CP-478; pp. 1/1-1/9). Neuilly-Sur-Seine, France: NATO--Advisory Group for Aerospace Research and Development.

Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37,* 32-64.

Endsley, M. R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors, 37,* 65-84.

Everitt, B. S. (1993). *Cluster Analysis*, 3rd ed. NY: Halsted Press.

Fracker, M. L. (1991). *Measures of situation awareness: Review and future directions* (Tech. Report AL-TR-1991-0128). Wright Patterson Air Force Base, OH: Air Force Systems Command.

Glaser, R., & Chi, M. T. H. (1988). Overview. In M. T. H. Chi, R. Glaser, and M. J. Farr (Eds.), *The nature of expertise* (pp. xv-xxviii). Hillsdale, NJ: Erlbaum.

Guzzo, R. A., & Salas, E. (Eds.) (1995). *Team effectiveness and decision-making in organizations.* San Francisco, CA: Jossey-Bass, Inc.

Guzzo, R. A., & Shea, G. P. (1992). Group performance and intergroup relations in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 3, 2nd ed., pp 269-313). Palo Alto, CA: Consulting Psychologists Press.

Hackman, R. A. (Ed.). (1990). *Groups that work (and those that don't): Creating conditions for effective team work.* San Francisco: Jossey-Bass.

Hinsz, V. B. (in press). Metacognition and mental models in groups: An illustration with metamemory of group recognition memory. To appear in E. Salas, S.M. Fiore, & J.A. Cannon-Bowers (Eds.), Team Cognition: Process and Performance at the Inter- and Intra-Individual Level. Washington, DC: American Psychological Association.

Hinsz, V. B. (1999). Group decision making with responses of a quantitative nature: The theory of social decision schemes for quantities. *Organizational Behavior and Human Decision Processes, 80,* 28-49.

Kiekel, P. A., Cooke, N. J., Foltz, P. W., Gorman, J., Martin, M. (2002). Some promising results of communication-based automatic measures of team cognition. *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting.* Santa Monica, CA: Human Factors and Ergonomics Society.

Kiekel, P. A., Cooke, N. J., Foltz, P. W., & Shope, S. M. (2001). Automating measurement of team cognition through analysis of communication data. In M. J. Smith, G. Salvendy, D. Harris, and R. J. Koubek (Eds*.), Usability Evaluation and Interface Design*, pp. 1382-1386, Mahwah, NJ: Lawrence Erlbaum Associates.

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor. *Journal of Management, 20,* 403-437.

Kraiger, K., & Wenzel. L. H. (1997). Conceptual development and empirical evaluation of measures of shared mental models as indicators of team effectiveness. In M. T. Brannick, E. Salas, & C. Prince (Eds.). *Performance Assessment and Measurement: Theory, Methods, & Applications* (pp. 63-84). Mahwah, NJ: Erlbaum.

Hutchins, E. (1991). The social organization of distributed cognition. In L. B. Resnick, J. M. Levine, S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 283-307). Washington, DC, USA: American Psychological Association.

Landauer, T. K., Foltz, P. W., and Laham, D. (1998). An introduction to Latent Semantic Analysis. *Discourse Processes, 25*, 259-284.

Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors, 42,* 242-271.

Levine, J. M., & Moreland, R. L. (1990). Progress in small group research. *Annual Review of Psychology, 41,* 585-634.

Orasanu, J. (1990). Shared mental models and crew decision making. (Tech. Rep. No. 46). Princeton, NJ: Princeton University, Cognitive Science Laboratory.

Prince, C., Stout, R. & Salas, R. (1994). *Team situational awareness preliminary lessons learned: Identification of skills, behaviors, design guidelines, and training strategies from interviews and research.* Technical Report, FAA and Naval Air Warfare Center, Training Systems Division.

Salas, E. & Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias, & J. D. Fletcher (Eds.), *Training & retraining: A handbook for business, industry, government, and the military* (pp. 312-335). New York: Macmillan Reference.

Salas, E. Dickinson, T. L., 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: Ablex.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors, 37,* 123-136.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *International Journal of Aviation Psychology, 1,* 45-57.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37,* 5-19.

Schvaneveldt, R. W. (1990). *Pathfinder associative networks: Studies in knowledge organization.* Norwood, NJ: Ablex.

Schvaneveldt, R. W., Durso, F. T., Goldsmith, T. E., Breen, T. J., Cooke, N. M., Tucker, R. G., & DeMaio, J. C. (1985). Measuring the structure of expertise. *International Journal of Man-Machine Studies, 23,* 699-728.

Steiner, I. D. (1972). *Group processes and productivity.* NY: Academic Press.

Stout, R., Cannon-Bowers, J. A., & Salas, E. (1996). The role of shared mental models in developing team situation awareness: Implications for training. *Training Research Journal, 2,* 85-116.

Taylor, R. M. & Selcon, S. J. (1994). Situation in mind: Theory, application, and measurement of situational awareness. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.), *Situational awareness in complex systems* (p. 69-77). Daytona Beach, FL: Embry -Riddle Aeronautical University Press.

Vidulich, M. A. (1989). The use of judgment matrices in subjective workload assessment: The subjective workload dominance (SWORD) technique. *In Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1406-1410). Santa Monica, CA: Human Factors and Ergonomics Society.

Wegner, D. M. (1986). Transactive Memory: A contemporary analysis of the group mind. In B. Mullen and G. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.

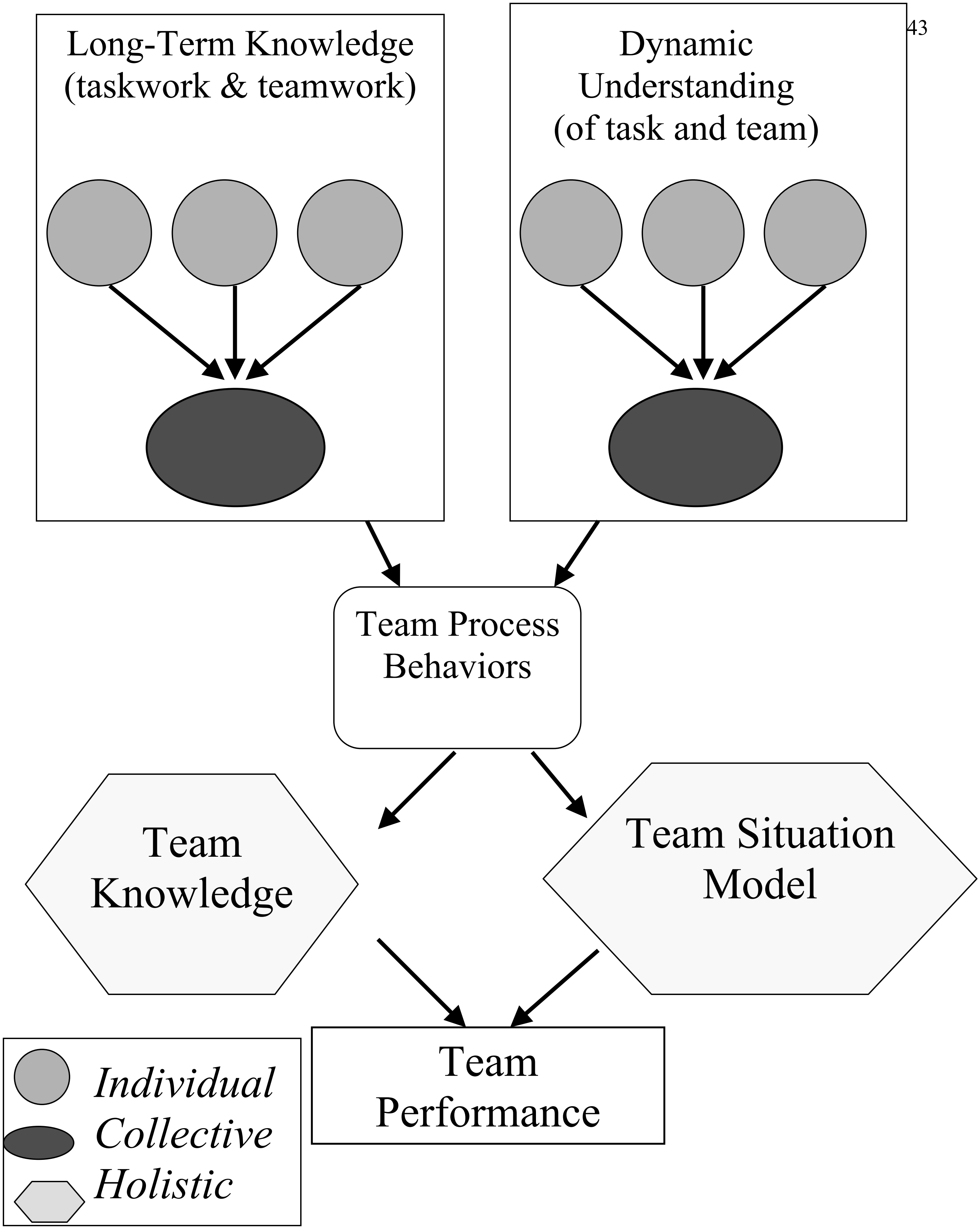
Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating Beliefs in Conversation. *Journal of Memory and Language, 31*, 183-194.

Table 1. Team cognition measurement needs.

1. Measures that target the holistic level, rather than the collective level, of team cognition.
2. Measures of team cognition suited to teams with different roles.
3. Methods for aggregating individual data to derive collective knowledge.
4. Measures of team knowledge that target the more dynamic and fleeting situation models.
5. Measures that target different types of team knowledge (e.g., strategic, declarative, procedural knowledge or taskwork vs. teamwork knowledge).
6. The extension of a broader range of knowledge elicitation methods to the problem of eliciting team cognition.
7. The streamlining of measurement methods to better automate them and embed them within the task context.
8. The validation of newly developed measures.

Figure Caption

Figure 1. Team knowledge framework.


Long-Term Knowledge
(taskwork & teamwork)
Dynamic
Understanding
(of task and team)
Team Process
Behaviors
Team
Knowledge
Team Situation
Model
Team
Performance
Individual
Collective
Holistic