



Australian Government
Department of Defence
Defence Science and
Technology Organisation

Measurement of the Behavioural, Cognitive, and Motivational Factors Underlying Team Performance

*Christopher Shanahan, Christopher Best, Melanie Finch and
Christopher Sutton*

Air Operations Division
Defence Science and Technology Organisation

DSTO-RR-0328

ABSTRACT

In the military, the use of teams to perform complex work functions is pervasive. Team performance is therefore critical, across a range of work contexts, for fulfilling operational requirements. To understand team performance, it is necessary to measure team performance. Valid and reliable measurement of the variables underlying effective team performance is fundamental to both team training and team training research, as it enables the determination of the strengths and weaknesses of teams, the design of targeted training programs, the provision of constructive feedback, and the evaluation of training effectiveness. This report presents, for the first time in a single resource, an overview of both the variables underlying effective team performance and the ways in which these variables have been operationalised and measured in the research literature.

RELEASE LIMITATION

Approved for public release

Published by

*Air Operations Division
DSTO Defence Science and Technology Organisation
506 Lorimer Street
Fishermans Bend Victoria 3207 Australia*

*Telephone: (03) 9626 7000
Fax: (03) 9626 7999*

*© Commonwealth of Australia 2007
AR-013-933
June 2007*

APPROVED FOR PUBLIC RELEASE

Measurement of the Behavioural, Cognitive, and Motivational Factors Underlying Team Performance

Executive Summary

In the military, as in other government organisations and industry, the use of teams to perform complex work functions is pervasive. The increasing reliance on teams has been driven by factors such as increasing workplace complexity, specialisation of labour, and organisational interdependence. Effective teams are seen to possess greater potential to manage work environments of this nature than are employees working independently (Cooke, Salas, Cannon-Bowers, & Stout, 2000; Salas, Sims, & Burke, 2005). Team performance is therefore critical across a range of work contexts for the fulfilment of operational requirements.

The ability of personnel to contribute as members of a team in a dynamic military environment can be enhanced through training strategies aimed at providing competencies which facilitate teamwork. An essential part of training is measurement; valid and reliable measurement of the variables underlying effective team performance is necessary in order to determine the strengths and weaknesses of a team, design targeted training programs, provide constructive feedback, and to evaluate the success of training.

The purpose of this report is to present, for the first time in a single resource, an overview of the manner in which the variables underlying effective team performance have been operationalised and measured in the research literature. The variables selected for inclusion in this report are consistent with a review of the past 50 years of research on small groups and teams by Kozlowski and Ilgen (2006), which identified the key team processes and emergent states that reliably influence team effectiveness – and which therefore offer leverage points that can be used to make teams more effective through deliberate intervention. According to Kozlowski and Ilgen (2006), the view of teams that emerges from this body of literature is one of complex, dynamic systems that exist in a context, develop over time, and evolve and adapt with the situational demands placed upon them.

The scope of this report is restricted to the measurement of factors that are applicable across a wide variety of domains. The measurement of task-specific teamwork inputs and outputs is not addressed. Although the variables and measures outlined in this report are relevant to many teams, they are geared toward ‘action’ teams (Kozlowski & Ilgen, 2006; Salas & Cannon-Bowers, 1997). Action teams are characterised by specialised member roles and responsibilities, and work environments that are dynamic, complex, stressful, and time-pressured. Examples of such teams include military command and control (C2), fire fighting, nuclear power process control, and surgical teams. In these teams effective communication, coordination, and decision making are crucial, and effective team training is vital.

This report uses the distinction between behavioural, cognitive, and motivational factors as an organising principle. It is divided into three main sections, each covering the measurement of one of these three kinds of teamwork factors. Where relevant, factors affecting the validity of the constructs and measurement techniques are discussed in an

effort to provide both direction on the nature of the techniques and the information required to make a judgment about their relative merit.

This report concludes that although it may seem attractive to simply base inferences about team performance on observations of the outcomes of team tasks, such an approach does not take into account the processes and emergent states which gave rise to those outcomes. Accordingly, outcome-based approaches are of little utility to researchers striving to understand teams, or to training specialists striving to achieve improvements in teamwork. In contrast, the measurement of the behavioural, cognitive, and motivational constructs discussed in this report enables a more direct examination of the true determinants of team task outcomes. We hope that the material presented here will allow both researchers and training specialists to make informed judgments about how to tailor measurement approaches to their specific needs.

References:

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

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7(3), 77-124.

Salas, E., & Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. A. Quinones & A. Ehrenstein (Eds), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 249-279). Washington, DC: American Psychological Association.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a “big five” in teamwork? *Small Group Research*, 36(5), 555-599.

Authors

Christopher Shanahan

Air Operations Division

Christopher Shanahan is a Research Scientist in the Crew Environments and Training (CET) branch of Air Operations Division. He is a forensic psychologist by training, and after completing his postgraduate studies he worked from 1999 to 2002 as a clinician in the NSW Department of Corrective Services' Violent Offender Treatment Program. Mr. Shanahan then commenced a PhD at the University of New South Wales, where he also served as Clinical Coordinator of the postgraduate program in forensic psychology. He joined DSTO in 2006, and his research interests include emotion and cognition, psychological assessment techniques, and training.

Christopher Best

Air Operations Division

Dr Christopher Best is a Research Scientist in the Crew Environments and Training (CET) branch of Air Operations Division. Dr Best holds a B.A.(Hons) in psychology (awarded in 1997) and a PhD in psychology (awarded in 2001). He was a member of the academic staff of the School of Psychology at Deakin University's Melbourne Campus for three years before joining DSTO in 2004. His research interests include human perception, cognition, teamwork, team performance measurement, and training.

Melanie Finch

Air Operations Division

Melanie Finch works in Crew Environments and Training within the Air Operations Division. Melanie has a B.Sc. (Hons) in psychology (awarded in 2005), and joined DSTO in 2006. Melanie's interests include modelling helicopter systems and crewing, human perception, team training and measurement, and data analysis.

Christopher Sutton

Air Operations Division

Christopher Sutton is a Research Scientist within Air Operations Division's Crew Environments and Training branch (CET). Christopher joined DSTO in 2006. He holds a B.Comm and a B.A.(Hons) in psychology (awarded in 2005) and is currently completing a Masters degree in organisational psychology.

Contents

1. INTRODUCTION	1
2. MEASURING BEHAVIOURAL FACTORS.....	4
2.1 The Team-Dimensional-Training Approach.....	5
2.2 The Anti-Air Teamwork Observation Measure.....	7
2.3 Teamwork Observer Rating Forms	8
2.4 Behavioural Observation in Aircrew Coordination Training.....	10
2.5 The TARGETS/EBAT Approach.....	13
2.6 Communication Behaviours	16
3. MEASURING COGNITIVE FACTORS: TEAM KNOWLEDGE.....	20
3.1 Team knowledge: Mental Models	22
3.1.1 Individual Mental Models.....	22
3.1.2 Team Mental Models	22
3.1.3 The Function and Utility of Team Mental Models.....	23
3.1.4 The Content of Mental Models	23
3.1.5 Multiple Mental Models	24
3.2 The Meaning of Shared: Homogenous and Heterogeneous Knowledge.....	24
3.2.1 Transactive Memory	25
3.3 Team Mental Models and Team Performance: Empirical Evidence	25
3.4 The Measurement of Team Knowledge	27
3.4.1 Observation	27
3.4.2 Interview	28
3.4.3 Process Tracing and Verbal Protocols	30
3.4.4 Conceptual Methods	30
3.4.5 Issues in Eliciting and Measuring Knowledge	31
3.5 Team Knowledge Metrics	32
4. MEASURING COGNITIVE FACTORS: SITUATION AWARENESS.....	33
4.1 Team Situation Awareness	33
4.2 The Measurement of Individual Situation Awareness	35
4.2.1 Indirect Measures of Situation Awareness	35
4.2.1.1 Performance Measures	35
4.2.1.2 Physiological Measures	36
4.2.2 Direct Measures of Situation Awareness	37
4.2.2.1 Subjective Measures	37
4.2.2.2 Objective Measures	38
4.2.3 The Global Implicit Measure	39
4.3 The Measurement of Team Situation Awareness.....	40
4.3.1 Examples of Aggregation Measures of Team Situation Awareness	41
4.3.2 Measures Specific to Team Situation Awareness.....	42
4.3.3 Team Mutual Awareness.....	45
4.4 Conceptual Issues with Team Situation Awareness.....	46

5. MEASURING MOTIVATIONAL FACTORS	47
5.1 Collective Efficacy	47
5.1.1 Antecedents and Effects of Collective Efficacy	47
5.1.2 Measurement of Collective Efficacy.....	48
5.2 Team Cohesion.....	49
5.2.1 A Model of Military Team Cohesion	50
5.2.2 Carron's Hierarchical Model of Cohesion	51
5.3 Trust.....	52
5.3.1 Antecedents and Effects of Trust.....	53
5.3.2 Measurement of Trust.....	54
5.4 Applications and Conclusions	54
 6. CONCLUSION.....	 54
 7. REFERENCES.....	 55
 APPENDIX A: SCALES FOR THE MEASUREMENT OF MOTIVATIONAL FACTORS	 69

1. Introduction

In the military, as in other government organisations and industry, the use of teams to perform complex work functions is pervasive. The increasing reliance on teams has been driven by factors such as increasing workplace complexity, specialisation of labour, and organisational interdependence. Effective teams are seen to possess greater potential to manage work environments of this type than are employees working independently (Cooke, Salas, Cannon-Bowers, & Stout, 2000; Salas, Sims, & Burke, 2005). Team performance is therefore critical, across a range of work contexts, for the fulfilment of operational requirements.

In the military context the changing nature of work environments is best exemplified by the rise of the concept of network-centric warfare, which offers the potential for enhanced operational speed, adaptability, and efficiency. The promise of networked operations is likely to depend, however, on the ability of individuals in military teams to develop a common understanding of intent, develop an accurate knowledge of their own roles and responsibilities as well as the roles and responsibilities of others, build and maintain situation awareness, and communicate and collaborate effectively (Best, Sauvage, & Skinner, in press). Importantly, however, team competencies such as these cannot be assumed. Assigning a team to perform a particular task, or simply building the infrastructure to support a collaborative style of operations, does not in itself guarantee the desired outcome: it is not enough to simply ask individuals to work together and expect that they will know how best to do so (Rousseau, Aubé, & Savoie, 2006).

The ability of personnel to contribute as members of a team in a dynamic military environment can be enhanced through the use of training strategies aimed at providing competencies which facilitate teamwork. An essential part of team training and training research is measurement. However, it is not sufficient to simply measure the outcomes of team taskwork, as outcome measures are of little diagnostic value. Valid and reliable measurement of the variables *underlying* effective team performance is necessary to determine the strengths and weaknesses of teams, design targeted training programs, provide constructive feedback, and evaluate the success of training. The purpose of this report is to present, for the first time in a single resource, an overview and critique of the manner in which the variables underlying effective team performance have been operationalised and measured in the research literature.

In order to describe current approaches to team performance measurement, it is necessary to first identify the key variables underlying team performance. Unfortunately, the literature on team processes has long been characterised by a profusion of theoretical models, each identifying different key variables, or using different names to describe the same variables (Militello, Kyne, Klein, Getchell, & Thorsden, 1999). Although there have recently been several commendable attempts to lessen the confusion in the literature by integrating the existing disparate approaches (e.g., Marks, Mathieu, & Zaccaro, 2001; Militello et al., 1999; Rousseau et al., 2006; Salas et al., 2005), this report adopts a more pragmatic approach and focuses exclusively on those teamwork variables whose relation to team performance is supported by a substantial, systematic body of empirical research. These variables are displayed in

Table 1.1, where they are classified according to whether they are behavioural, cognitive, or motivational in nature.

The variables selected for inclusion in this report are consistent with a review of the past 50 years of research on small groups and teams by Kozlowski and Ilgen (2006), which identified the key team processes and emergent states that reliably influence team effectiveness – and which therefore offer leverage points that can be used to make teams more effective through deliberate intervention. According to Kozlowski and Ilgen (2006), the view of teams that emerges from this body of literature is one of complex, dynamic systems that exist in a context, develop over time, and evolve and adapt with the situational demands placed upon them. Their conceptual framework of team performance is outlined in Figure 1.1. In this framework, the team is embedded in a broader system context and dynamic task environment. The team is required to resolve problems and deal with challenges through a coordinated process that combines their cognitive, motivational/affective, and behavioural resources. A team is effective only when these team processes are aligned with environmentally-driven task demands. Team effectiveness, which Kozlowski and Ilgen regarded as synonymous with team outputs, reciprocally influences the environment in an ongoing cycle.

This conceptualisation of team effectiveness is based on the traditional Input-Process-Output (IPO) heuristic, but differs in that it recognises that many of the factors that mediate the translation of inputs to outputs are not behavioural processes, but emergent cognitive or motivational states (Ilgen, Hollenbeck, Johnson, & Jundt, 2005). These emergent states – such as situation awareness and team cohesion – characterise properties of the team that are dynamic in nature, and vary as a function of team context, inputs, processes, and outcomes (Marks et al., 2001). According to Marks et al. (2001), emergent states serve as both team inputs *and* outputs: they are the product of team processes and experiences, and then become inputs that influence the execution of subsequent teamwork processes. Regardless of their classification, they have the potential to explain variability in team performance and viability (Ilgen et al., 2005).

Table 1.1. Behavioural, cognitive, and motivational dimensions underlying team performance

Behavioural processes	Cognitive structures and processes	Motivational processes and states
Information exchange	Mental models	Collective efficacy
Communication	Transactive memory	Cohesion
Monitoring behaviour	Situation awareness	Trust
Supporting behaviour		
Leadership and initiative		
Coordination		
Cooperation		
Assertiveness		
Decision making		
Adaptability		

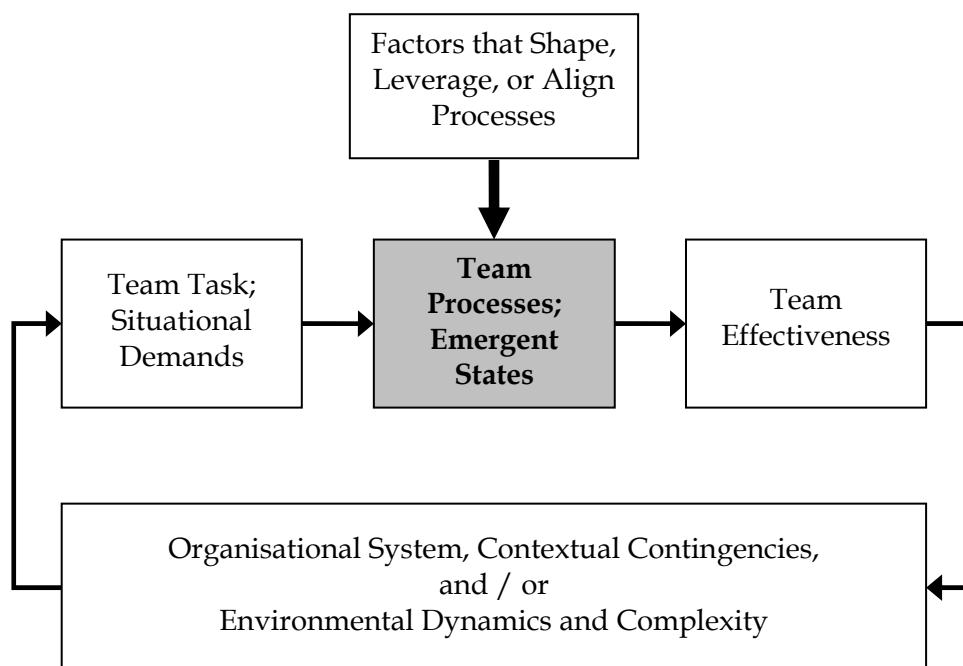


Figure 1.1. Conceptual framework of team effectiveness. From “Enhancing the effectiveness of work groups and teams” by W. J. Kozlowski and D. R. Ilgen, 2006, *Psychological Science in the Public Interest*, 7, p.79.

The scope of this report is restricted to the measurement of team processes and emergent states (represented by the shaded portion of Figure 1.1). The measurement of task-specific teamwork inputs and outputs is not addressed. Although the variables and measures outlined in this report are relevant to many teams and work domains, they are geared toward ‘action’ teams (Kozlowski & Ilgen, 2006; Salas & Cannon-Bowers, 1997). Action teams are characterised by specialised member roles and responsibilities, and work environments that are dynamic, complex, stressful, and time-pressured. Examples of action teams include military command and control (C2), fire fighting, nuclear power plant process control, and surgical teams. In these teams effective communication, coordination, and decision making are crucial, and effective team training is vital.

This report uses the distinction between behavioural, cognitive, and motivational factors as an organising principle. It is divided into three main sections, each covering the measurement of one of these three kinds of teamwork factors. Where relevant, factors affecting the validity of the constructs and measurement techniques are discussed in an effort to provide both direction on the nature of the techniques and information to assist in making judgments about their relative merit. We anticipate that this report will be useful in guiding the design of measurement strategies for research and training events in which there is a focus on teamwork issues. We begin consideration of measurement approaches with a discussion of measurement of the behavioural factors underlying team performance.

2. Measuring Behavioural Factors

The development of a measure of teamwork behaviour generally proceeds in three stages. The first is the identification and operationalisation of categories of teamwork behaviours, which are often labelled 'dimensions' of teamwork. The second is the codification of the goals, tasks, and behaviours of a specific team in terms of these dimensions of teamwork. While this can be accomplished in many ways, an example of a formalised method for describing teamwork skills in support of behavioural measurement is the hierarchical task analysis for teams (HTAT) approach, championed most notably by Annett (e.g., 1997; Annett, Cunningham, & Mathias-Jones, 2000). This approach is centred on a hierarchical decomposition of the team's goals and is aimed at providing a detailed description of the teamwork behaviours that take place within a particular context. The third stage of measure development involves the construction of behavioural checklists, rating scales, or other performance indicators relevant to the identified dimensions, goals, tasks, and behaviours. The teamwork research literature is replete with lists of teamwork dimensions and their associated behaviours, and the lack of a single, unified conception of teamwork and its measurement is much lamented (e.g., Marks et al., 2001; Militello et al., 1999). However, significant commonalities can be seen among the various approaches. For example, Cannon-Bowers, Tannenbaum, Salas, & Volpe (1995) reviewed the literature on teamwork and identified a relatively small number of skill dimensions and associated subskills that have been consistently employed across a range of research contexts (see Table 2.1, overleaf).

Given the sheer number of different ways in which teamwork has been operationalised, and the similarities that often exist between the various approaches, only the most prominent frameworks will be described here. This section of the report will discuss the application of these frameworks to specific work domains, and the means by which measures of the relevant behaviours have been developed. The aim is to provide the reader with an overview of *what* has been measured, and *how* these measurements have been applied to the investigation of behavioural processes in teamwork.

The use of observer ratings is common in the field of team behaviour measurement, and therefore much of this section of the report is concerned with different methods for structuring and capturing observer ratings. First, two approaches to team behaviour measurement developed by Smith-Jentsch and her colleagues will be described. While both are based on observer ratings, one provides qualitative data, the other quantitative data. Second, several alternative methods for capturing quantitative data using observer ratings will be discussed. These alternative methods are useful in that they either offer a different definition of the dimensions of teamwork, or employ a different rating format. Third, the Targeted Acceptable Responses to Generated Events or Tasks (TARGETS) technique is described. This approach is valuable to consider because it represents an attempt to reduce the subjectivity of observer ratings. Finally, approaches to the measurement of one particularly important class of teamwork behaviours – communication behaviours – is described.

Table 2.1. Teamwork skill dimensions and subskills

Skill Dimension	Subskills
Adaptability	Flexibility Compensatory Behaviour
Shared Situational Awareness	Dynamic Reallocation of Functions Situational Awareness Shared Problem-Model Development
Performance Monitoring/Feedback	Intra-member Feedback Mutual Performance Monitoring
Leadership/Team Management	Task Structuring Mission Analysis Motivation of Others
Interpersonal Relations	Conflict Resolution Cooperation (Interpersonal) Assertiveness Morale Building (Behavioural Reinforcement)
Coordination	Boundary Spanning Task Organisation Task Interaction
Communication	Timing and Activity Pacing Information Exchange Consulting with Others
Decision Making	Problem Assessment Problems Solving Planning Metacognitive Behaviour Implimentation (Jurisdiction)

Note. Adapted from “Defining team competencies and establishing team training requirements,” by J. A. Cannon-Bowers, E. Salas, S. I. Tannenbaum, C. E. and Volpe, 1995, in R. Guzzo, E. Salas & Associates (Eds.), *Team Effectiveness and Decision Making in Organizations*, pp. 344-346. San Francisco: Jossey-Bass.

2.1 The Team-Dimensional-Training Approach

Smith-Jentsch, Zeisig, Acton, and McPherson (1998) and Smith-Jentsch, Johnston, & Payne (1998) analysed performance data from US Navy combat information centre (CIC) teams, and identified four distinct behavioural dimensions underlying teamwork. Effective teams were seen to be comprised of individuals who exhibited behaviours related to: (1) information exchange, (2) communication, (3) supporting behaviour, and (4) initiative and leadership. As described by those authors, information exchange involves knowing what to pass to whom and when, communication focuses on how information is delivered, supporting behaviour involves actions taken by team members to compensate for one another, and initiative and leadership focuses on behaviours that provide direction for the team. Each of these dimensions was further defined in terms of component behaviours, as displayed below in Table 2.2.

Table 2.2. Smith-Jentsch et al.'s (1998) conceptualisation of teamwork dimensions and their associated component behaviours

Teamwork Dimension	Component Behaviours
Information Exchange	• Utilising information from all available sources • Passing information to the appropriate persons before having to be asked • Providing situation updates that summarize the big picture
Communication	• Using proper phraseology • Ensuring that reports are complete • Avoiding excess/unnecessary chatter (brevity) • Using a clear, audible tone of voice
Supporting Behaviour	• Monitoring for errors and taking action to correct those errors when they occur • Requesting and offering backup or assistance to adjust workload among team members
Leadership and Initiative	• Offering guidance or suggestions to others • Stating clear and appropriate priorities

In developing these dimensions, Smith-Jentsch and her colleagues extracted what they regarded as measurable, team-level constructs from an initial pool of candidate dimensions. In doing so, they discarded dimensions and associated behaviours which (1) required assessors to infer a state of mind on the part of the team members, or (2) could be argued to reflect individual, rather than team components of performance. It is interesting to note, however, that Smith-Jentsch et al.'s definition of 'Communication' (i.e., using proper phraseology, ensuring that reports are complete, avoiding excess/unnecessary chatter, and using a clear, audible tone of voice) could be argued to be largely reflective of individual performance, rather than team performance.

The measurement technique that Smith-Jentsch, Zeisig et al. (1998) developed in order to capture performance along these dimensions formed part of an approach to managing team-training events called Team Dimensional Training (TDT). This technique primarily provides qualitative data, and is geared specifically towards the provision of rich feedback on team performance. The measurement technique described by these authors involves three steps. First, instructors seek out and record details of both positive and negative behavioural examples of an assigned TDT dimension¹, using the component behaviours to guide their

¹ It is important to note that, due to assessor workload, the authors recommended one instructor per dimension.

observations. This is done on a worksheet that provides an exercise timeline, including a description of exercise scenario events, with space for notes to be written beside each. Then, after the exercise has been completed, instructors meet to summarise their observations in preparation for an after-action review (AAR) with exercise participants. During this meeting, instructors select positive and negative examples of each of the component behaviours in order to provide discussion points in the AAR. These are recorded on simple summary sheets which provide two boxes for text to be entered next to each of the component behaviours; one for positive examples and one for negative examples. Finally, on the basis of the AAR, instructors diagnose the team's (1) current strengths, and (2) goals for future improvement on each teamwork dimension. These are recorded on a 'Teamwork Summary' sheet and used as inputs to the team's planning for the next training exercise.

Given the deliberately qualitative nature of this measurement approach, it is well suited to the provision of feedback during training. However, it is less well suited for research purposes, or for readiness management (i.e., tracking the competency status of organisational units over time), as it is difficult to compare performance across teams or measurement occasions. These latter purposes demand a more quantitative approach, such as those described in the following sections.

2.2 The Anti-Air Teamwork Observation Measure

Smith-Jentsch and her colleagues (Smith-Jentsch et al., 1998) also applied the four dimensions of teamwork used in the TDT approach in the development of a quantitative instrument for assessing teamwork called the Anti-Air Teamwork Observation Measure (ATOM). The process of capturing teamwork behaviours using the ATOM consists of three steps. The first step is identical to the first stage of data collection in the TDT approach described above. In the second step of the ATOM procedure, following the exercise the instructors review their notes and assign ratings against each of the component behaviours corresponding to the four dimensions of teamwork (see Table 2.2). According to Smith-Jentsch et al. (1998) these ratings are assigned on a three-point scale according to how frequently *negative* instances of the component behaviours were observed during the exercise: either zero, between one and three times, or more than three times. For example, in relation to the communication dimension, instructors assess how frequently (1) the team engaged in the use of improper phraseology, (2) reports provided to team mates were incomplete, (3) team members engaged in excess or unnecessary chatter, and (4) team members spoke without using a clear, audible tone of voice. In the third step of the ATOM procedure, instructors provide overall ratings of the team's performance on each of the four teamwork dimensions. These ratings are provided on a five-point, behaviourally-anchored rating scale (BARS). The lower end is anchored using negative descriptors, and the upper end is anchored using positive descriptors. Examples of the two rating scales used in the ATOM procedure are provided in Figure 2.1 (overleaf).

A		B	
Communication Behaviour Rating		Overall Communication Rating	
Improper Phraseology	0 1-3 >3	12345 Consistently demonstrated improper phraseology, excess chatter, incomplete reports, or inaudible communication No instances of improper phraseology, excess chatter, incomplete reports, or inaudible communication	
Inaudible Communication	0 1-3 >3		
Excess Chatter	0 1-3 >3		
Incomplete Reports	0 1-3 >3		

Figure 2.1. Examples of the rating scales used as part of the ATOM procedure. Panel A shows the frequency-based scale used to rate component behaviours. Panel B shows the BARS used to capture overall ratings of each teamwork dimension. Adapted from "Measuring team-related expertise in complex environments," by K. A. Smith-Jentsch, J. H. Johnston, and S. C. Payne, 1998, in J. Cannon-Bowers and E. Salas (Eds.), *Making Decisions Under Stress: Implications for individual and team training*, pp. 68-69. Washington, DC: APA Press.

Smith-Jentsch et al. (1998) reported that the ATOM approach displayed predictive validity in respect to team outcomes, was useful in diagnosing shortcomings in team performance, and was sensitive to the effects of team training interventions. A particularly salient finding from their evaluation was that composite teamwork dimension scores, formed by summing ratings of component behaviours, performed better than overall dimension ratings in diagnosing team performance. Unfortunately the authors did not discuss whether this was due to the specificity of the items (component behaviours versus overall dimensions), or to the difference in the nature of the scales used to rate performance (frequencies versus evaluative judgments).

2.3 Teamwork Observer Rating Forms

In addition to describing the four-dimensional ATOM instrument outlined above, Smith-Jentsch et al. (1998) described a process by which that tool and the framework associated with it were developed from earlier theoretical conceptualisations of teamwork. This framework was also used as the basis for the development of a separate series of teamwork rating scales by Elliot and Eileen Entin and their colleagues (e.g., Serfaty, Entin, & Johnston, 1998; Entin & Serfaty, 1999; Entin & Entin, 2001; MacMillan, Entin, Hess, & Paley, 2004; MacMillan, Paley, Entin, & Entin, 2004). This research group have developed a number of teamwork rating scales, each tailored to a specific research setting. However, these scales have typically been based on the generic teamwork dimensions of communication, monitoring, feedback, backup, coordination, and team orientation. While most descriptions of the use of these scales make reference to all six dimensions, some applications have involved as few as four. Unfortunately, descriptions of the scale appearing in the literature have not typically specified exactly which teamwork dimensions were included. A version of the scale obtained from the authors in late 2005 (in which the 'feedback' dimension is omitted) is displayed in Table 2.3.

Table 2.3. *Generic teamwork dimensions underpinning the measurement approaches of Entin, Entin and colleagues*

Teamwork Dimension	Description
Communication Behaviour	Communication involves the exchange of information between two or more team members in the prescribed manner and using proper terminology. Often the purpose of communication is to clarify or acknowledge the receipt of information.
Monitoring Behaviour	Monitoring refers to observing the activities and performance of other team members. It implies that team members are individually competent and that they may subsequently provide feedback and backup behaviour.
Backup Behaviour	Backup behaviour involves assisting the performance of other team members. This implies that team members have an understanding of other members' tasks. It also implies that members are willing to give and seek assistance.
Coordination Behaviour	Coordination refers to team members executing their activities in a timely and integrated manner. It implies that the performance of some team members influences the performance of other team members. This may involve an exchange of information that subsequently influences another member's performance.
Team Orientation	Team orientation refers to the commitment team members exhibit to working together. It implies that they place the goals and interest of the team ahead of their personal goals. It also refers to the trust each team member has in the other team members, team pride, and esprit de corps.

Note. From the form used in the Adaptive Architectures for Command and Control (A2C2) research program. Personal communication, 2005.

The teamwork behaviour measures based on these dimensions were specifically designed to be independent of the task domain and mission objectives associated with any particular team or scenario (MacMillan et al., 2004). Although they have typically been used to capture expert-observer ratings of team performance, they have also been modified for use as self report measures in which team members report on their own perceptions of team processes (MacMillan et al., 2004). In the typical application of the scales, observers are asked to take notes on teamwork behaviours during scenario execution, then to rate items associated with each dimension on 7-point BARS at the conclusion of the mission. As illustrated in Figure 2.2, the lower and upper end points of the scales for each item are anchored by examples of behaviours indicative of poor and good team processes.

Communication Behaviour	
1.	To what extent were error caused by inadequate team communication?
	1 2 3 4 5 6 7
7	Communication within the team was always effective and never responsible for errors or inefficient performance
1	Communication was wholly inadequate and resulted in inefficient performance and many of the errors made by the team
2.	To what extent did team members provide relevant information to other team members, in a proactive way, without having to be asked for it?
	1 2 3 4 5 6 7
7	Team members always provided relevant information to others without being asked
1	Team members never provided relevant information to others unless specifically asked

Figure 2.2. Examples of the BARS items used in the teamwork behaviour rating scales developed by Entin, Entin and their colleagues. Personal communication, 2005.

An advantage of this measurement approach is that the context-independent wording of the items makes them broadly applicable, with little effort required to tailor measurement to particular contexts. A shortcoming of the approach is that there is not a great deal of information in the open literature about the psychometric properties of these scales. Entin and Serfaty (1995) reported inter-rater reliabilities of .79 or higher for the scale, and Serfaty et al. (1998) reported internal consistency so high that it could in fact be indicative of some degree of item redundancy (coefficient alpha = .98). While the authors often cite support for the construct validity of the scale from factor analytic studies, Serfaty et al. (1998) reported that principal components analysis of responses on one version of the scale extracted just two factors (i.e., far fewer than the purported six dimensions), which accounted for nearly 90% of the variance in assigned scores. Each of the items on the scale loaded at 0.7 or above on one of those two factors. The first factor was related to the giving and receiving of assistance, while the second factor was related to communication and anticipation. Serfaty et al. did not provide specific information about which items loaded on which factor. These findings raise doubts about the construct validity of the scale, and caution should be exercised when interpreting the outcome of an assessment based on the scores for each different dimension.

2.4 Behavioural Observation in Aircrew Coordination Training

Brannick, Roach, and Salas (1993) developed and evaluated a behavioural measure of team performance that was used in the context of a low-fidelity (i.e., PC-based) air combat team

task. This work was based on earlier developments by the US Navy Training Systems Centre (Morgan, Glickman, Woodard, Blaiwes, & Salas, 1986; and Glickman et al., 1987; both cited in Brannick et al., 1993). The generic teamwork dimensions employed by Brannick et al. are displayed in Table 2.4. Although little specific information was provided about the nature of the items or the rating scale used, this approach is worth noting as Brannick et al. reported evidence suggesting that their measures possessed reasonable internal consistency, were predictive of team outcomes, and were sensitive to differences between teams in terms of member familiarity and previous experience with the flight simulation task.

Table 2.4. Brannick et al.'s (1993) teamwork dimensions and associated definitions

Teamwork Dimension	Definition
Technical Coordination	Acting properly in sequence and providing useful information.
Interpersonal Cooperation	Reflects the quality of team member interchanges, such as whether they argued or gave encouragement.
Giving Suggestions	Not explicitly defined
Accepting Suggestions	Not explicitly defined

Note: Brannick et al. (1993) also included measurement of 'Team Spirit' in their measurement approach. As this construct is most appropriately conceptualised as motivational, rather than behavioural, it is omitted here.

A modified version of Brannick et al.'s (1993) rating scale was used by Volpe, Cannon-Bowers, Salas & Spector (1996) to measure teamwork processes in an investigation of the effects of cross-training on team performance. Volpe et al. changed the name of the 'Interpersonal Cooperation' dimension to 'Interpersonal Coordination,' and generated a new dimension called Cross-Monitoring. No explicit description of this new dimension was given, but it may have been intended to account for behaviours similar to Brannick et al.'s (1993) dimensions of 'Giving Suggestions' and 'Accepting Suggestions.' Volpe et al. summed ratings on items associated with each of these dimensions to produce an overall teamwork process score. Scores on individual dimensions were not analysed or reported.

Brannick, Prince, Prince, and Salas (1995) gave an account of the development of teamwork behaviour measurement instruments in the context of aviation crew coordination training for the US Navy. While the authors explicitly noted that some of their items were directly relevant to that particular domain, they also asserted that the dimensions they identified should be "important to effective functioning across many types of teams" (p.643). The six dimensions and their definitions are presented in Table 2.5. Although Brannick et al. (1995) did not provide specific examples of the items contained in their measurement instrument, it is nevertheless important to take this approach into account in order to understand (1) the dimensions of behaviour that these authors considered to be relevant for teamwork in general, and (2) the measurement techniques they employed in capturing teamwork behaviours.

Table 2.5. Brannick et al.'s (1995) teamwork dimensions and associated definitions

Teamwork Dimension	Definition
Assertiveness	Willingness to make decisions and act on them, to defend decisions, and to admit ignorance and ask questions.
Decision making/Mission analysis	Using sound judgment to select the best course of action based on available information. Seeking information and allocating and monitoring resources.
Adaptability/Flexibility	Altering a course of action in the face of changing conditions, appropriate change of action, and maintaining constructive behaviour under pressure.
Situational Awareness	Identifying problems, maintaining an accurate perception of the external environment, and detecting situations that require corrective action.
Leadership	Directing and coordinating the activities of other crew members and monitoring crew performance.
Communication	Clearly and accurately sending and acknowledging information, instructions, or commands.

These dimensions were used to construct two different kinds of measurement instruments. The first was a behavioural observation record consisting of the observable behaviours assigned to each of the six dimensions. This record took the form of a checklist in which the behaviours associated with each teamwork dimension were not rated, but rather scored dichotomously (YES/NO) depending upon whether or not the instructor observed an occurrence of the behaviour during mission execution. While the authors did not state the nature of these items explicitly, it is likely, given the way that this scale was constructed, that the items were all positive examples of teamwork behaviours (i.e., the more that were observed, the higher the team's score for the mission). This dichotomous approach to measuring team behaviours is similar to the TARGETS approach (i.e., Targeted Acceptable Responses to Generated Events or Tasks; e.g., Fowlkes, Lane, Salas, Franz, & Oser, 1994) which will be discussed in detail below. The second measurement instrument, called the rating guide, consisted of a series of BARS items which provided instructors with the opportunity to assign a score from 1 (poor performance) to 5 (good performance) on each of the teamwork dimensions. As is characteristic of BARS items, descriptive anchors were provided for each scale point, based on behaviours considered indicative of performance expected at that particular level.

Brannick et al. (1995) used the multi-trait multi-method (MTMM) technique to examine the validity of their instruments, and found them to possess good convergent and discriminant validity. There was also evidence that their behavioural process measures were related to

team outcomes, and useful for discriminating between novices and experts. However, the results were less positive in terms of test-retest reliability, with different patterns of results for the same teams across different mission scenarios. The authors made the point that characteristics of the mission scenario may play a significant role in determining team performance, and that multiple measurement occasions may be required in order to build a true picture of teamwork.

It is worth noting that some of the dimensions of teamwork that were used by Brannick et al. (1995) are not consistent with the rationale employed by Smith-Jentsch and her colleagues in constructing the TDT/ ATOM teamwork dimensions. As noted above, Smith-Jentsch and her colleagues discarded dimensions and associated behaviours which (1) required assessors to infer a state of mind on the part of the team members, or (2) could be argued to reflect individual, rather than team components of performance. Brannick et al's (1995) dimensions of Situation Awareness and Decision Making/Mission Analysis would most likely be excluded according to these criteria (the measurement of team situation awareness is discussed further in Section 4 of this report).

2.5 The TARGETS/EBAT Approach

A widely applied method for measuring teamwork behaviours, particularly in the context of team training, is the Targeted Acceptable Responses to Generated Events or Tasks (TARGETS) technique. This measurement technique has most often been associated with an event-based approach to team training (EBAT; e.g., Fowlkes, Dwyer, Oser & Salas, 1998); indeed, it is difficult to describe the TARGETS approach without first describing EBAT. Generally, event-based approaches to the design of training scenarios have in common a focus on presenting to trainees, through the deliberate introduction of specific types of scenario events, the opportunity to demonstrate (and therefore learn) specific classes of competencies. Training designers seeking to apply event-based approaches typically do so by systematically identifying (1) desirable training outcomes, (2) contextualised knowledge, skills, and attitudes related to those outcomes, (3) events that stimulate trainees to engage in behaviours directly related to such knowledge, skills, and attitudes, and (4) a method for supporting behavioural-observation measures of these competencies. An important aspect of the event-based approach is that the process of defining these components of training should include a focus on the identification of explicit linkages *between* training objectives, targeted competencies, training scenario events, and measurement. During the training exercise, each event is designed so that it creates the requirement to undertake behaviours related to identified competencies. Measurements are made by assessors, based on the successful (or otherwise) demonstration of those behaviours by the team under observation.

9-line provided to fighters		Event A: Heavy Enemy Artillery	
- IP		Fighters informed of situation	
- Heading (IP to target)		Fighters told to orbit at safe dist	
- Distance (IP to target)		Fighter plan time considered	
- Target description		Informal ACA established/re-established	
- Type, mark, laser code, target line		Alternate ACA considered	
- Friendly location		ACA	
- Egress direction		- Protects fighters	
- Remarks: SEAD, ACA clearly described, etc		- Allows fighter manoeuvrability	
- TOT/TTT		- FSO ensures indirect fires conform to ACA	

Comments: _____

1 = Observed/Performed Satisfactorily
0 = Omitted/Failed to Perform Satisfactorily
x = No Opportunity to Perform/Not Required

Figure 2.3. An example of part of a TARGETS checklist showing behaviours which should be observed in the context of a close air support training exercise. Adapted from "Team performance measurement in distributed environments," by D. Dwyer, J. Fowlkes, R. L. Oser, E. Salas, and N. E. Lane, 1997, in M. T. Brannick, E. Salas, and C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications*, p. 148. Mahwah, NJ: Lawrence Erlbaum Associates.

The TARGETS measurement technique is commonly used to capture team performance in training events structured according to an EBAT. As described by Fowlkes et al. (1994) and Dwyer, Fowlkes, Oser, Salas, and Lane (1997), the TARGETS approach builds on the EBAT structure by providing assessors with opportunities to observe and record specific behaviours related to each event within the exercise scenario. This has typically involved the development of a time-ordered checklist which includes an inventory of scenario events and specific, observable behaviours that the team should undertake in response to them. Assessors use the checklist to indicate (in a YES/NO format) whether appropriate team behaviours were observed in response to each training scenario event. Scores are based simply on the number (rather than the thoroughness, efficiency, accuracy, or timeliness) of behaviours that are observed. An example of a partial TARGETS checklist, adapted from Dwyer et al. (1997) is presented in Figure 2.3. This checklist was developed in the context of a close air support training mission, and covers the behaviours associated with the 'routine' event of passing of the 9-line brief (left-hand checklist), as well as the 'nonroutine' event of dealing with heavier than usual enemy artillery fire (right-hand checklist).

There are many examples of the application of the TARGETS approach (e.g., Dwyer et al., 1997; Fowlkes et al., 1994; Fowlkes et al., 1998; Salas et al., 1999; Proctor, Panko, & Donovan, 2004; Stout, Salas, & Fowlkes, 1997). While this body of research follows the same basic method (outlined above) in structuring measurements, some differences can be seen in the way that the content of the TARGETS checklists are developed. In particular, differences exist in the attention paid to generic dimensions of teamwork during development. Some applications of the TARGETS approach have involved checklists structured around team goals and associated tasks, with little explicit consideration given to generic teamwork dimensions. An example of this approach is that presented by Dwyer et al. (1997). These authors reported that their TARGETS checklist behaviours were based on doctrine, task analyses, standard operating procedures, SME input, and other documented tactics, techniques, and procedures. However, they did not explicitly state whether, or which, teamwork dimensions they took into account.

Other applications of the TARGETS approach have focused more explicitly on generic dimensions of teamwork as an organising principle, while also taking into consideration many of the task-specific sources of information listed above. For example, Stout et al. (1997) developed their TARGETS checklist through a process which involved making “skill dimensions and the generic behaviours related to these skill dimensions ... context specific for the scenario that was performed” (p.175). This resulted in a list of specific, effective teamwork behaviours that were associated with team goals and scenario events as well as underlying generic teamwork dimensions. Stout et al. specifically focused on the observation of behaviours related to three of the teamwork skill dimensions (and related subskills) identified by Cannon-Bowers et al. (1995) in their review of the team competencies literature: Communication, Assertiveness, and Situational Awareness. The reason Stout et al. focused on these dimensions was that they believed them to be most relevant to their particular research setting.

Some of the claimed benefits of the TARGETS approach are that the clearly-defined structure of the observation forms simplifies the rating task, that less training and expertise are required than traditional rating techniques because of the YES/NO format of the ratings, and that the close linkages developed between team goals, scenario events, behaviours, and scoring enhances the objectivity, specificity, and diagnosticity of ratings (Dwyer et al., 1997). Acceptable inter-rater reliability and internal consistency have also been reported for some TARGETS checklists (e.g., Fowlkes et al., 1994).

However, the approach also has shortcomings. One of these is that the tight coupling between the TARGETS checklist, exercise scenario events, and very specific sequences of behaviours makes it difficult to deal with unforeseen events – or with the flexible, adaptive behaviours that may arise in response to these events. This aspect of the approach also precludes the reuse of TARGETS checklists and increases the amount of effort required to develop them, as they are tied specifically to particular scenarios. Another drawback is that the dichotomous scoring technique (cited by the developers of the approach as an advance over rating scale techniques) affords the assessor no ability to assign ratings based on the *quality* of behavioural responses (e.g., their thoroughness, efficiency, accuracy, or timeliness). This has the potential to significantly reduce the richness of the information that is recorded via this measurement technique, which may in turn negatively impact upon data analysis, performance feedback, and ultimately the utility of research or training outcomes.

2.6 Communication Behaviours

It is clear from the above discussion that the way in which teamwork behaviours have been conceptualised and operationalised varies widely. One constant theme in the research literature, however, is the recognition of communication as critical to team performance. As highlighted above, it is debatable whether some aspects of communication behaviour are best conceptualised as individual-level or team-level skills. For example, Smith-Jentsch et al.'s (1998) definition of communication, framed in terms of *how* information is delivered, explicitly relates to aspects of individual linguistic ability such as the use of proper phraseology, and speaking in a manner that can be easily understood. Other aspects of communication behaviour relate more explicitly to team processes. For example, passing information to team mates in a proactive manner, acknowledging the receipt of information, providing situation updates to establish and maintain situation awareness, asking questions to clarify ambiguities, and speaking up when a problem is noticed are all aspects of communication that fit more closely with the widely cited generic dimensions of teamwork behaviour.

The approaches that we have discussed so far have relied on observer ratings to establish whether, and how effectively, the team displays communication behaviours. In this section, alternative approaches to measuring communication behaviours are considered. First, classification strategies, based on frequency counts and the identification of different types of utterances, are described. Second, a more statistically-intensive, content analysis approach known as latent semantic analysis (LSA) will be discussed.

Although communication is almost universally regarded as an important teamwork process, there is generally not a direct relationship between the amount of communication data recorded during a training or research event and the usefulness of that data for analysis or feedback. In particular, when communications among team members are captured at a very detailed level (e.g., as verbatim transcripts), it is difficult and time consuming to develop meaningful, quantitative measures that describe the nature of the communications (Entin & Entin, 2001). However, quantitative records comprised of simple frequency counts of communications events are not particularly useful either. This is because, by themselves, such frequencies provide little information about the quality of teamwork processes – sometimes effective teams engage in many communication behaviours, sometimes they engage in very few (e.g., Entin & Serfaty, 1999). Classification strategies, on the other hand, are aimed at capturing communication data at an intermediate level of detail that is based primarily on frequencies, but which captures some information regarding the semantic content of team utterances.

In an investigation of the effects of cross-training on teamwork processes, Volpe et al. (1996) obtained observational ratings of various teamwork dimensions in addition to classifying communication events according to five categories: Requesting Information, Volunteering Information (Observation, Command, or Suggestion), Indicating Agreement or Compliance, Task-irrelevant Remarks, and Acknowledgements. Volpe et al. hypothesised that cross-trained teams would communicate more appropriately than teams that did not receive cross-training. In this case, more appropriate communication was operationalised as more information volunteered, more acknowledgements and agreements offered, fewer requests for information, and fewer task-irrelevant comments. Volpe et al. (1996) found significant effects for the cross-training manipulation in terms of the behavioural ratings of teamwork processes,

indicating better processes in cross-trained teams. However, they found only mixed support for cross-training in terms of communication variables. Cross-trained teams did volunteer more information than those who did not receive cross-training, but no significant effects were found for most communication variables – and on task-irrelevant remarks, the cross-trained teams actually showed poorer performance.

Serfaty et al. (1998) also made use of communications variables in their investigation of team adaptation and coordination training (TACT) for Navy CIC teams. In this study communications events were classified not only in terms of their content, but also in terms of the team members involved in the message (i.e., sender and receiver). This was done by constructing matrices in which the columns represented different senders and receivers (e.g., tactical action officer to tactical information coordinator) and the rows represented different categories of communications (e.g., request for information). Each cell therefore represented a particular combination of team members and communication type (e.g., request for information from tactical action officer to tactical information coordinator). Coders placed tally marks within each cell of the matrix to indicate the frequency with which each event took place during the exercise scenario. Serfaty et al. divided communications events into Requests and Transfers – equivalent to Volpe et al.'s (1996) use of requesting and volunteering – as well as Acknowledgements. The former two categories were further divided according to whether communication events related to Information, Action and Task, or Problem Solving and Planning; this scheme therefore yielded a total of seven categories of communication events.

From this data Serfaty et al. (1998) extracted a number of variables with which they characterised information flow within the CIC team. A representation of the pattern of sending and receiving information within the team was obtained by examining the number of communications sent by each team member (expressed as a percentage of the total number of communications events recorded), as well as by examining the destination of communications. Communication destination referred to whether messages were sent downward (tactical action officer to other team members), upward (team members to tactical action officer), laterally (within the team, excluding the tactical action officer), or outward (from the team to external agencies). Communication data was also used to examine anticipatory behaviour within the team. The ratio of transfers to requests (called the 'Anticipation Ratio') was calculated for several categories of communication behaviours. In interpreting such a ratio, values greater than one indicate that team members 'pushed' information more frequently than they 'pulled' information. This implies anticipation of the information needs of other team members (McMillan et al., 2004). While Serfaty et al. only reported detailed results for a global anticipation ratio (i.e., the ratio of total transfers to total requests), the authors claimed that both the global ratios and the more specific ratios (e.g., the anticipation ratio for upward information) revealed an overall increase in anticipation as a result of TACT training.

Entin and Entin (2001) reported an updated version of the communications matrices used by Serfaty et al. (1998), which preserved the logic of the measurement strategy but utilised slightly different categories of communications. Entin and Entin (2001) maintained the high-level differentiation between requests, transfers, and acknowledgements. However, they divided the former two categories differently. Requests were divided into Information, Action/Task, and Coordination. Transfers were divided into Information, Will Perform

Action/Task, and Will Coordinate. Some of the measures extracted from this measurement technique were described by these authors in a table, which is reproduced as Table 2.6.

Fischer, McDonnell and Orasanu (2007) used three different classification strategies in their study of team communication during a simulated search and rescue task. These researchers classified the task-related content of communication events in terms of whether they involved (1) information sharing (e.g., regarding tasks, logistics, location, or movement), (2) problem solving (e.g., setting goals, making plans or predictions), (3) meta-cognition (e.g., assessing or reporting on own or team progress), or (4) team coordination (e.g., directing the action of others or stating one's intentions). In addition to studying the task-related content of communication events, Fischer et al. also examined their affective content and the interactive patterns observed between the members of their four-person teams during task execution. For the former, communication events were coded as either positive (e.g., humour, praise, empathy), neutral (e.g., politeness, appeasement), or negative (e.g., irony, insult, defensiveness) in affective valence. For the latter, Fischer et al. coded the team members' responses to each others utterances as constituting (1) acknowledgements, (2) disagreements, (3) elaborations, (4) answers to questions, or (5) missing responses.

Table 2.6. Measures of verbal communication commonly derived from the communications matrix technique

Measure	Description
Overall Rate	Total number of communications per minute
Information requests	Number of requests for information per minute
Information transfers	Number of transmissions of information per minute
Action requests	Number of requests for an action per minute
Action transfers	Number of statements of actions (to be) undertaken per minute
Coordination requests	Number of requests to coordinate an action per minute
Coordination transfers	Number of agreements to coordinate an action per minute
Acknowledgements	Number of non-substantive acknowledgements of receipt of communication (e.g., 'ok' to acknowledge receipt of information) per minute
Overall anticipation ratio	All communication transfers divided by all communication requests
Information anticipation ratio	Information transfers divided by information requests
Action anticipation ratio	Action transfers divided by action requests

Note: In order to compare data from scenarios of different lengths, Entin and Entin (2001) recommend that measures be calculated per unit time, typically per minute, rather than as raw totals. Adapted from "Measures for evaluation of team processes and performance in experiments and exercises," by E. Entin, & E. Entin, 2001, *Proceedings of the 6th International Command and Control Research and Technology Symposium*, p. 8. Annapolis, MD.

Fischer et al.'s (2007) analysis of the content of communications demonstrated that these strategies were useful for discriminating between high- and low-performing teams. Analysis of the task-related content revealed a difference in information sharing; a larger proportion of the utterances of high-performing teams were dedicated to information sharing than was the case for low performing teams. Teams also differed in terms of the affective content of their communications, with members of high-performing teams showing more humour, praise and empathy, while members of low-performing teams showed more insults and defensiveness. Finally, teams differed in terms of their interactive patterns, with high-performing teams elaborating more than their low-performing counterparts, while missing responses were more frequent for low- than high-performing teams.

Estimates vary of the training that is required to ensure proficiency and inter-rater reliability in coding communication. Entin and Entin (2001) recommended that observers be trained in the use of the communications matrix method for around 25 to 30 hours in order to ensure an acceptable level of proficiency in coding. Evidence indicating that a degree of proficiency in coding communications could be achieved relatively quickly was obtained from research conducted during Exercise Pacific Link 1 (Crane et al., 2006), a coalition distributed mission training exercise held between the Defence Science and Technology Organisation (DSTO) in Australia and the Air Force Research Laboratory (AFRL) in the United States. In Pacific Link 1, two raters independently coded communications between two pairs of simulated fighter aircraft (two Royal Australian Air Force F/A-18s and two US Air Force F-16s) and an air battle manager during simulated air combat missions. In this exercise there was reasonable agreement between the coders from the outset, despite observers having only a few hours practice on communications standards in the air combat domain. Nevertheless, agreement did increase over the course of the three-day exercise, indicating that more practice beforehand may have resulted in better outcomes.

Cooke, Salas, Keikel, & Bell (2004) provided an overview of a relatively sophisticated, statistical approach to the analysis of communication behaviours known as latent semantic analysis (LSA). LSA is a statistical technique that can be used to code communications content in order to represent and analyse semantic information underlying the use of terms within a particular domain (Landauer, Foltz, & Laham, 1998). The technique begins by establishing the co-occurrence of words within a large body of text (e.g., a transcript from an exercise scenario). This is calculated by defining a matrix of the frequency of co-occurrence of each word with each other word. In the example given by Cooke et al., an instance of co-occurrence is defined as the two words occurring in the same paragraph. The technique then proceeds in a similar way to principle components analysis or factor analysis: through the extraction of latent dimensions underlying the data in the co-occurrence matrix. It is these underlying dimensions which are purported to measure elements of meaning in the text. However, as in other statistical techniques of this kind, the user can define the number of dimensions to be extracted, which introduces a degree of subjectivity to the technique. Cooke et al. describe their use of LSA to examine variables such as the similarity of a given team's dialogue to that of other teams whose performance is known, and the internal consistency of the team's dialogue.

Gorman, Foltz, Kiekel, Martin, and Cooke (2003) described several measures that can be extracted as part of LSA for characterising team communication behaviours. Among these were communication density (CD) and lag coherence (LC). The measurement of CD is

motivated by the idea that the effectiveness of team communication is related to the transfer of information in the most *concise* manner possible. CD is therefore calculated as the ratio of meaningfulness (or task-relevance; as determined via the outcomes of the LSA procedure) to the number of words used (Gorman et al. note that this is analogous to the way that average velocity is calculated in kinematics, with meaningfulness being analogous to distance covered and number of words used being analogous to time taken). The measurement of LC is used to capture topic shifting across utterances. This is motivated by the idea that there should be some relationship between sequential utterances during team tasks. If sequential utterances address completely different concepts, then the team is hardly communicating at all. If sequential utterances address exactly the same concepts throughout an entire mission, then team members are simply repeating information that has been previously conveyed to them.

Because of the novel measures that are being derived from the method, LSA holds promise as a means of quantifying aspects of communication which may be related to team performance. However, there are challenges which must be overcome if it is to become more widely used. The relationship between LSA-based measures and more standard team performance measures, as well as the power of LSA-based measures to predict team outcomes, are yet to be established across a variety of contexts. Furthermore, the application of LSA requires team communications to be transcribed into text; this represents a logistical problem since transcription is an extremely labour intensive task for human coders, and voice recognition software is, at the present time, less than completely reliable.

3. Measuring Cognitive Factors: Team Knowledge

In most environments, we depend on coordinated cognitive interactions with others to function successfully (Levine, Resnick, & Higgins, 1994). This is particularly so in the case of teamwork, which involves people working together to achieve something that is beyond the capabilities of individuals working alone (Marks et al., 2001). Establishing an understanding of the cognitive requirements associated with this interdependency, referred to broadly as *team cognition* (Cooke et al., 2000), is widely acknowledged to be critical to understanding the mechanisms underlying effective teamwork, and to developing interventions geared at improving team performance (Cannon-Bowers & Salas, 2001).

The pivotal role that team cognition is seen to play in teamwork has generated a considerable body of theoretical and applied research (for reviews of the team cognition literature see Klimoski & Mohammed, 1994; and Mohammed & Dumville, 2001). Unfortunately, in this literature team cognition has been interpreted in diverse ways, and few researchers have clearly defined or explicitly described what they mean by the central construct (Klimoski & Mohammed, 1994). This lack of conceptual clarity has resulted in conflicting and occasionally contradictory findings, and has limited the value of this research to organisations, industry, and the military.

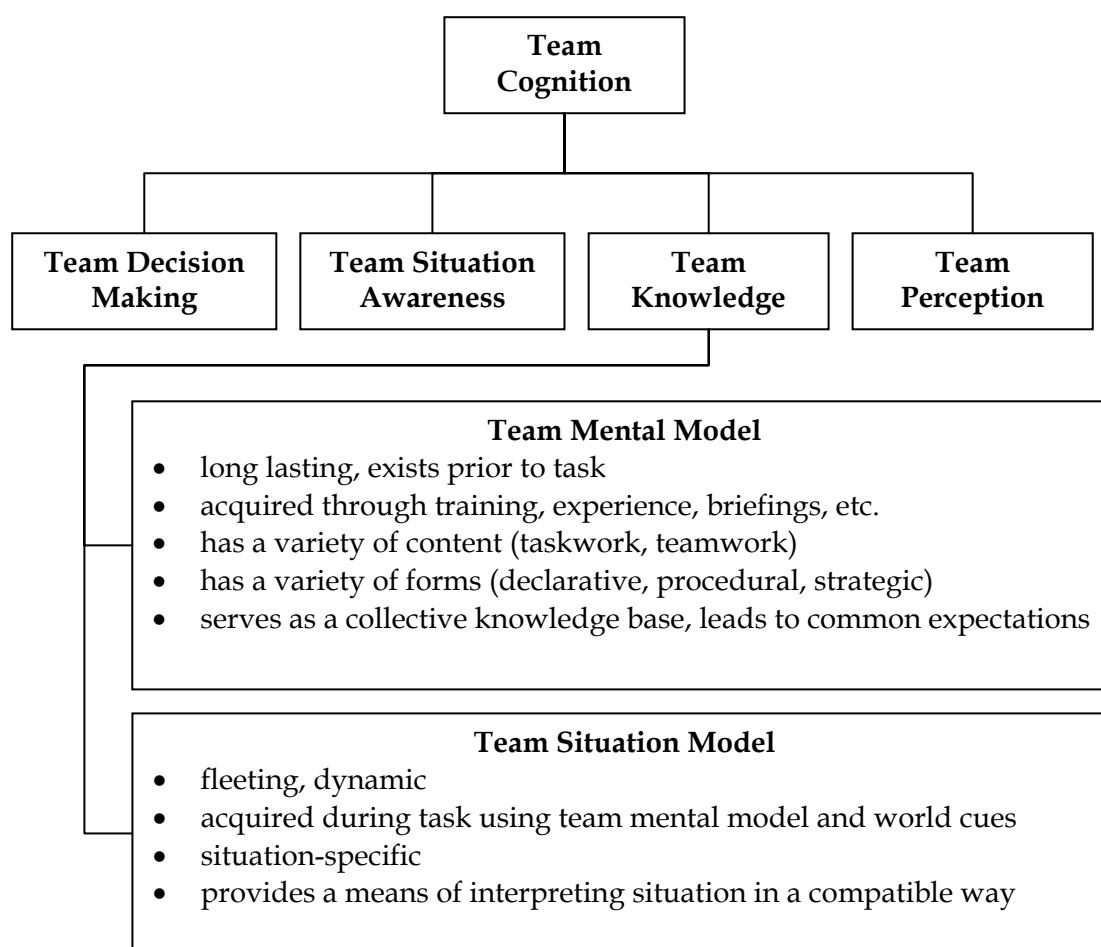


Figure 3.1. Team Cognition Framework. Adapted from “Measuring team knowledge” by N. J. Cooke, E. Salas, J. A. Cannon-Bowers, and R. J. Stout, 2000, *Human Factors*, 42, p. 153.

This report adopts the theoretical framework of team cognition proposed by Cooke et al. (2000), outlined in Figure 3.1. Consistent with the vast bulk of the applied literature on team cognition, the focus of this report is on the subset of team cognition referred to as team knowledge². According to Cooke et al. team knowledge is comprised of the team mental model and the team situation model. The team mental model refers to the task- and team-related knowledge held by team members, and will be discussed first. The team situation model, which is more commonly conceptualised as situation awareness (SA), refers to the team’s collective understanding of the current situation. The measurement of team SA will be discussed in detail in Section 4.

² In this theoretical framework Cooke et al. (2000) make no assumptions about the form of representation of team knowledge, and use of the term ‘knowledge’ is neutral with regard to the accuracy or completeness of that information.

3.1 Team knowledge: Mental Models

This section of the report is divided into two parts. The first part discusses (1) the nature of individual and team mental models, (2) the function and utility of team mental models, (3) the notion of 'shared' mental models and the distribution of knowledge within teams, and (4) the empirical evidence concerning the relationship between team mental models and performance. The second part provides an overview of the major techniques used to elicit and represent team knowledge, and summarises the metrics commonly used to quantify this information for the purpose of examining the similarity or accuracy of individual and team mental models.

3.1.1 Individual Mental Models

Mental models refer to a general class of cognitive constructs invoked to explain how knowledge and information are represented in the mind (Klimoski & Mohammed, 1994). Outside of the applied literature they are more formally and precisely referred to as scripts (Schank & Abelson, 1977), schemas (Anderson, 1976), or prototypes (Rosch & Mervis, 1975). In the simplest sense, mental models are representations of what an individual knows about something at a given time (Suppes, Pavel, & Falmagne, 1994). However, they are more than simply 'knowledge' in general; they are concerned with the organisation of, and relations between, pieces of information about some concept (Mohammed & Dumville, 2001). The abstract or general knowledge they represent also holds across many particular instances of that concept (Fiske & Taylor, 1991). The function of mental models is to affect the interpretation of related information, in particular the construal of ambiguous information and the values that are assumed for unavailable information (Smith, 1998). In applied terms, they serve the purpose of providing a framework for describing, explaining, and predicting system states (Rouse & Morris, 1986).

Mental models are not complete, veridical representations of the environment; rather, they are representations of how people conceive of that environment (Cohen & Murphy, 1984). It is neither necessary nor desirable for mental models to completely represent the external domain under consideration³: as their utility lies in speed and efficiency, mental models preserve important detail and ignore that which is deemed unimportant (Cohen & Murphy, 1984; Schneider, 1991).

3.1.2 Team Mental Models

Prior to defining team mental models, it is worth pointing out that the term 'shared mental model,' commonly used in the applied literature, implies that a single mental model can be distributed amongst several team members. This is misleading. Mental models, as they are traditionally construed, are strictly individual-level constructs. It is vital to recognise that the team mental model is a qualitatively different construct to the individual mental model. Klimoski & Mohammed (1994) and Cooke et al. (2000) in fact explicitly use the term 'team'

³ Cohen and Murray (1984) point out that mental models are in fact twice removed from the environment: they are theoretical representations of mental representations that in turn represent the external environment.

mental model in favour of 'shared' mental model due to the potential for confusion arising from its disparate meanings: share can mean either to hold in common, or to apportion. The relevance of the term 'shared' to team mental models is discussed later in this section.

The most commonly accepted definition of the team mental model is "an organised understanding of relevant knowledge that is shared by team members" (Mohammed & Dumville, 2001, p. 89). This parsimonious definition, which implies a commonality of knowledge, belies the complexity of the construct. First, although usually measured at the individual level, the team mental model is not simply an aggregation of individual mental models. Rather, it is viewed as an emergent characteristic of the group, resulting from an interaction between each individual's cognition and team process behaviours, in particular communication (Cooke, Salas, Kiekel, & Bell, 2004; Klimoski & Mohammed, 1994). Second, and in a related sense, team mental models only exist to the extent that they are apprehended, at some level of awareness, by team members. For example, having every team member independently recognise a problem is not sufficient for identification of the problem at the team level (Klimoski & Mohammed, 1994). Cooke et al. (2000) give the analogy of a cockpit crew where the pilot is aware of a threat at Point A, and the co-pilot is aware of a threat at Point B. Without effective communication between the crew, the sum of team knowledge does not represent the sum of the knowledge of each individual crew member. A team mental model requires both a common understanding of relevant knowledge (i.e., actuarial similarity), and a common awareness that each member of the team possesses this knowledge (Fiore & Schooler, 2004).

3.1.3 The Function and Utility of Team Mental Models

Team mental models purportedly serve orienting and coordinating functions. They allow individual team members to anticipate the needs of other team members, and to predict the behaviour both of other individuals within the team and of the group as a whole. Shared knowledge is also thought to enable team members to interpret cues in a similar manner, allocate resources appropriately, and to make compatible decisions (Cannon-Bowers & Salas, 2001; Klimoski & Mohammed, 1994).

Measurement of team knowledge is seen as necessary not only to understand the mechanisms driving effective team performance, but to design knowledge-intensive applications such as team-training programs and decision making technologies (Cooke et al., 2000). In addition, once the content and structure of team knowledge associated with effective performance has been established, this information can be employed to assess the current state of a team's knowledge in order to (1) gauge readiness to perform, (2) obtain an estimate of likely effectiveness, (3) diagnose problematic aspects of team performance, (4) identify training needs, and (5) evaluate the success of training interventions (Cannon-Bowers & Salas, 2001; Cooke et al., 2000).

3.1.4 The Content of Mental Models

Team mental models can include representations of tasks, equipment, working relationships, and situations; they can be declarative, procedural, or strategic in nature (Mohammed & Dumville, 2001). Cannon-Bowers and Salas (2001) describe four broad categories of team

knowledge: task-specific knowledge, task-related knowledge, knowledge of team-mates, and attitudes and beliefs.

Task-specific knowledge is more widely known as taskwork knowledge, and refers to the specific procedures, sequences, actions, and strategies necessary to perform a task. Task-related (or teamwork) knowledge refers to generic knowledge about teamwork processes, such as communication and supporting behaviour. Team mate knowledge concerns the roles, preferences, strengths, weaknesses, and tendencies of the team members. The final category of team knowledge concerns the general attitudes and beliefs held by team members; these attitudes and beliefs are neither task-specific nor task-related, but may influence team states such as motivation, cohesion, and consensus (Cannon-Bowers & Salas, 2001).

3.1.5 Multiple Mental Models

Klimoski and Mohammed (1994) caution against thinking about “a team mental model,” and stress that there are likely to be multiple mental models co-existing among team members at a given point in time. For example, a member of an Air Battle Management (ABM) team will have a mental model (or more accurately hierarchically organised mental models or *prototypes*) of the various hostile aircraft they may encounter: the prototype of a high fast flyer would include concepts such as altitude above 45,000 feet, speed greater than Mach 1, requires fast response and high priority, associated with a higher closure rate and extended weapon engagement zone. They will also have, among others, mental models of (1) how to operate their radar and communications equipment, (2) the role of the other members of the team, (3) the manner in which they are expected to interact with their team members, and (4) how a specific type of mission normally unfolds. Consequently, it is meaningless to speak broadly about ‘team’ or ‘shared’ mental models without providing a specific description of the mental model(s) in question.

3.2 The Meaning of Shared: Homogenous and Heterogeneous Knowledge

In the context of teamwork, research into mental models has been driven by the broad assumption that a shared understanding of the task, team, equipment, and situation is associated with greater team effectiveness and improved team performance (Mohammed & Dumville, 2001). Although the term ‘shared’ can encompass having in common and apportioning, the team mental model is typically operationalised as the degree of overlap among team members’ mental models⁴.

However, the idea that effective teams possess shared knowledge, in the sense of common or similar knowledge, is overly simplistic, and does not take into account the fact that team members are usually required to possess unique knowledge or skills (Mohammed & Dumville, 2001). A division of labour and a differentiation of roles and responsibilities is in fact a central principle in accepted definitions of teamwork (Salas, Dickinson, Converse, & Tannenbaum, 1992). Accordingly, some common or overlapping knowledge, in combination with some distributed or complementary knowledge, is more consistent with the accepted notion of a team (Cooke et al., 2000).

⁴ In this sense, ‘common’ is a more appropriate and accurate descriptor than ‘shared.’

Mohammed and Dumville (2001) point out that common or overlapping knowledge in teams with distinct roles may be inefficient. Wellens (1993), for example, suggests that an optimal group situation model is achieved when enough overlap occurs to maintain group coordination, while allowing enough division to maximise coverage of the environment. Mohammed and Dumville (2001) argue that instead of assuming that all team members need to have common knowledge in all domains, more attention should be directed towards specifying the domains and conditions under which common (homogenous) and distributed (heterogeneous) knowledge will either aid or hinder team performance. Furthermore, they argue that if overlapping knowledge is required, it is important to specify which team roles need to share what information: some team knowledge may need to be common to all team members, some knowledge may need to overlap among sub-groups of team members, and some knowledge may be unique to individual team members.

The relevance of the various interpretations of 'shared' therefore depends on the heterogeneity of the team. Mohammed and Dumville (2001) state that, for heterogeneous teams, performance is likely to be highest when there is a team mental model with primarily distributed taskwork knowledge, primarily overlapping teamwork knowledge, and a balance of diversity and consensus in belief structures; the optimal configuration of sharing across content domains will depend upon the nature of the team and the task.

3.2.1 Transactive Memory

The argument that knowledge is often heterogeneous within a team has much in common with the concept of transactive memory (Wegner, 1987). Transactive memory refers to a shared system for encoding, storing, and retrieving information (Wegner, Erber, & Raymond, 1991) that serves the purpose of distributing cognitive responsibilities so as to increase the cognitive efficiency of the group (Levine et al., 1994). In such a system, each member of the team learns in a general way what other team members know in detail; individual team member knowledge is specialised, but fashioned into a differentiated collective memory that is useful to the group (Mohammed & Dumville, 2001; Wegner et al., 1991). The key to developing an effective transactive memory involves the communication and updating of information each team member has about the areas of other team members' knowledge (Wegner et al., 1991). In that sense, a transactive memory system is composed of the memory systems of the individuals in the group *and* the communication processes that link these systems. A number of studies have found transactive memory accuracy to be positively related to group performance (Austin, 2003; Lewis, 2003), and Lewis (2003) suggests that a transactive memory system, in the form of "cooperatively divergent expertise," is an important accompaniment to knowledge that is shared between team members.

3.3 Team Mental Models and Team Performance: Empirical Evidence

The empirical literature on team mental models amounts to considerably less than the theoretical literature. Although both direct and indirect relationships have been established between various operationalisations of team mental models and team performance (Rentsch & Woehr, 2004), the findings are conflicting and occasionally contradictory.

Effective planning (Stout, Cannon-Bowers, Salas, & Milanovich, 1999) and cross-training (Marks, Sabella, Burke, & Zaccaro, 2002) have been associated with higher taskwork mental model similarity and better team performance. These performance gains were either associated with, or mediated by, an improvement in team processes such as communication, coordination, and backup behaviour. However, these findings were only partially supported by Mathieu, Heffner, Goodwin, Cannon-Bowers, & Salas (2005) who found that although teamwork mental model similarity led to improved performance through better team processes (strategy formation and coordination, cooperation, and communication), taskwork mental model similarity showed no relationship to team performance.

In a different vein, Edwards, Day, Arthur and Bell (2006) found mental model accuracy to be a much stronger predictor of team performance than mental model similarity, which is consistent with findings by Mathieu et al. (2005) that the quality of mental models mediates the relationship between mental model similarity and team performance. In contrast, Lim and Klein (2006) demonstrated that, while both taskwork and teamwork mental model similarity and accuracy were related to improved performance in infantry combat teams, the accuracy of the team's mental models did not moderate the relationship between teamwork mental model similarity and team performance.

The inconsistent and occasionally contradictory findings, particularly in relation to taskwork team mental models, can perhaps be explained by the fact that researchers have not considered the heterogeneity of knowledge within the teams they are examining. A notable exception is Cooke, Kiekel, Salas, Stout, Bowers, and Cannon-Bowers (2003), who found that the best predictors of team performance were high taskwork positional accuracy and *low* taskwork team knowledge similarity. In other words, team performance was associated with a division of knowledge among different roles rather than a similarity of knowledge among team members.

These findings highlight the need to specify precisely the conditions under which 'shared' knowledge might affect team performance outcomes (Cannon-Bowers & Salas, 2001). Espinosa, Lerch, & Kraut (2004) propose, for example, that the degree to which common mental models can improve coordination and performance depends on the nature of the dependencies present in a task, which in turn are influenced by a variety of task, team, and contextual variables. According to Espinosa et al. effective teams manage dependencies using a combination of explicit (e.g., oral and written communication, programming mechanisms such as schedules, plans, and procedures) and implicit (enabled by shared knowledge) coordination mechanisms and processes. If explicit coordination mechanisms are available, implicit coordination mechanisms may not be necessary, or may not have a strong influence on performance. However, when explicit coordination is difficult (e.g., as a result of excessive workload or time pressure) implicit coordination may be critical to team functioning, as it enables members to predict the information and resource requirements of their team mates.

Generally speaking, it is important to keep in mind that the relationship between team mental models and team performance is not straightforward. Klimoski & Mohammed (1994) summarise by sounding the cautionary note that team mental models are not necessarily functional; they can in fact also be dysfunctional.

3.4 The Measurement of Team Knowledge

Arguably the most critical aspect of measuring team knowledge is eliciting that knowledge in the first place. Therefore, the techniques that have been applied to knowledge elicitation will be discussed in detail here. Cooke et al. (2000) describe two broad approaches to the measurement of team knowledge: collective and holistic. The collective approach involves eliciting knowledge at the individual level, and then aggregating this data across team members to represent team knowledge. The holistic approach involves eliciting knowledge at the team level, and targets knowledge that results from the application of team processes to the knowledge of individual team members; it is described as *effective* knowledge that guides team action (Cooke et al., 2004). Cooke et al. (2000) argue that, while both forms of team knowledge are valid, team knowledge measured in the collective manner underestimates the importance of team processes in determining what the team knows.

The measurement techniques described in this section of the report fall exclusively under the umbrella of the collective approach. We have chosen this focus partly because these are the techniques currently employed in the empirical literature, and partly because there are no established holistic measurement methods⁵. There are four general approaches that can be used to elicit team knowledge: observation, interview, process tracing, and conceptual methods.

3.4.1 Observation

Although not traditionally considered a means of eliciting knowledge, observations (in written, audio, or video form) made in natural settings have the potential to provide a continuous source of data with minimal intrusion on task performance (Cooke et al., 2000). Observations are useful in terms of obtaining a global impression of the domain under consideration, and identifying factors that need to be considered during later, more targeted knowledge elicitation (Cooke, 1999). Langan-Fox, Code, & Langfield-Smith (2000) suggest that observation techniques need not necessarily be passive; the research can be actively involved in the process through questioning of the participant at regular (or relevant) points. The obvious disadvantage of observational techniques is that it is difficult to infer the nature of the individual or team mental model driving team behaviour; inductive reasoning on the part of the experimenter is therefore required (Gordon, Schmieder, & Gill, 1993). Furthermore, although observations provide a rich source of data, the amount of information generated can become unwieldy.

⁵ The difficulty with operationalising holistic knowledge has long been recognised, and Levine et al. (1994) acknowledge that it is typically not measured directly but instead inferred from successful coordination of activity by group members. Cooke et al. (2004) suggest interviewing the team as a whole when eliciting team knowledge, as this measurement may also capture the process behaviours of the team. However, the demand characteristics of this approach may induce process behaviours that differ from actual task-related process behaviours, and may in fact result in a false representation of team knowledge. A more promising approach suggested by Cooke et al. (2004) is to treat the communication within the team as a 'think aloud' protocol, and to analyse the flow and content of this data to obtain a holistic measure of team cognition (see the discussion on communication in Section 2.6 of this report).

3.4.2 Interview

The most direct way to elicit knowledge is to conduct an interview (Cooke, 1999). Unstructured interviews are an effective and flexible exploratory tool, and can be used to generate hypotheses which can be investigated later using more structured, focused, and systematic interviewing techniques. Examples of these more focused techniques, discussed below, include the Critical Decision Method (Klein, Calderwood, & MacGregor, 1989), conceptual graph analysis (Graesser & Clark, 1985), and conceptual mapping (Hoffman & Lintern, 2006). Once an adequate representation of a domain has been established, more straightforward interview techniques such as surveys, self-report questionnaires, and even multiple-choice tests can be employed to assess the current state of a team's knowledge.

The Critical Decision Method is a retrospective knowledge elicitation technique organised around a specific incident from the participant's own experience. The participant is guided in the recall and recounting of the incident, and then answers a series of progressively more specific queries (termed *cognitive probes*) intended to produce a comprehensive and detailed picture of the incident, and in particular the decision making strategies of the participant (Hoffman, Crandall, & Shadbolt, 2006; Hoffman & Lintern, 2006; Schraagen, 2006). A detailed description and application of this technique, which is typically employed in the creation of training material, is presented in Hoffman, Crandall, and Shadbolt (2006).

Graesser and Clark's (1985) knowledge acquisition and representation system was developed in order to represent both generic and specific knowledge structures. Generic knowledge structures refer to general conceptualisations such as schemas, scripts, and stereotypes (i.e., mental models); specific knowledge structures are cognitive representations that relate to a specific experience at a particular time and place (i.e., situation models). This procedure employs a question probe methodology (see Table 3.1) to elicit information from participants, and is expressly designed to uncover knowledge that is either implicit or extremely well learned⁶. The elicited knowledge is then represented in the form of a conceptual graph structure, which consists of a set of labelled statement nodes that are interrelated by labelled, directed arcs. Construction of a conceptual graph involves three major steps: (1) segmenting information into statement nodes (essentially a cognitive representation of a single state, event, process, or action), (2) assigning statement nodes to node categories, and (3) interrelating nodes via labelled directed arcs. A succinct description of the question probe methodology, and brief overview of conceptual graph analysis, is provided in Gordon et al. (1993). Conceptual graph construction is described in more detail in Graesser and Goodman (1985a and 1985b).

⁶ Hoffman and Lintern (2006) argue that there is no evidence that there exists such a thing as knowledge that cannot be articulated; the issue is whether the knowledge elicitation procedure provides a framework that is sufficient to support the expert in articulating what they know.

Table 3.1. A selection of generic question probes used to elicit knowledge for the construction of conceptual graphs

State or Event
What happens before _____?
What happens after _____?
What are the consequences of _____ occurring?
What causes or enables _____?
Goal
What happens before having the goal of _____?
What happens after having the goal of _____?
How is the goal of _____ attained?
What state or event initiates the goal of _____?
What is the outcome of _____?
Action
How do you _____?
What do you do before you _____?
What prevents you from being able to _____?
What do you do after _____?
What states or events cause or enable you to _____?
Why do you _____?
What are the consequences of _____?
What happens if you do not _____?

Note. Adapted from "Knowledge acquisition for instructional system design," by S. E Gordon, K. A. Schmieder, and R. T. Gill, 1993, *Human Factors*, 35, p. 469.

Concept mapping is similar to, and has its basis in, conceptual graph analysis. As the name implies, concept maps are diagrams which display concepts and the relationships among concepts. Hoffman and Lintern (2006) recommend the software suite CmapTools (<http://ihmc.us>), which employs a procedure that combines knowledge elicitation and knowledge representation – the concept map is created online in a collaborative process between the researcher and the participant.

The Cockpit Management Attitudes Questionnaire (Helmreich, Foushee, Benson, & Russini, 1986) provides an example of how a survey technique can be applied to the collection of information on knowledge structures, specifically with regard to team roles and responsibilities. The CMAQ was formed as a way to assess the benefits of resource management training. Examples of the items on the questionnaire are 'The captain should take control and fly the aircraft in emergency and non-standard situations' and 'There are no circumstances (except total incapacitation) where the first officer should assume control of the aircraft'. Helmreich et al. found that responses on this questionnaire could be used to distinguish experienced pilots from inexperienced pilots.

3.4.3 Process Tracing and Verbal Protocols

Process-tracing techniques involve collecting data concurrently with task performance, typically using a procedure in which the actor is asked to 'think aloud' and give verbal expression to their thoughts while engaged in the task (Ericsson & Simon, 1993). Less often, nonverbal data (including keystrokes, actions, facial expressions, eye fixations, and other behaviours) are used to trace cognitive processes and infer underlying knowledge (Cooke et al., 2000).

In verbal protocol analysis, it is important that the participant is instructed to verbalise their thoughts in a manner that minimises the potential for reactivity (Chi, 2006). Reactivity refers to the situation in which the act of generating a report changes the cognitive processes that mediate the observed performance (for a comprehensive description of protocol analysis methodology and applications see Ericsson & Simon, 1993; a cursory description is provided in Ericsson, 2006). As the verbal reports are elicited they are recorded, and then transcribed to enable protocol analysis.

Protocol analysis is concerned with assigning quantifiable measures of content and structure to the raw data, and involves developing and applying a coding scheme to capture the critical content. Coding schemes form the basis of quantifying knowledge, and are typically developed using a combined a priori and inductive approach (Miles & Huberman, 1994). Existing domain knowledge (e.g., from occupational specification documents, lists of core competencies) is used to provide a conceptual framework and provisional list of codes; this provisional list is then refined to fit events, and the relation between events, frequently referred to in the verbal protocol. The level of granularity in the coding system is an important consideration: the narrower the focus, the more detailed the representation of knowledge. This coding scheme is then applied to identified units (single events, processes, or actions) in the verbal protocol.

The most significant practical disadvantage of verbal protocol analysis is that it is an extremely labour intensive process, even when dedicated software (e.g., N6) is used to facilitate content analysis. The most significant theoretical criticism of protocol analysis is that the experimental situation may encourage individuals to construct a rationale for their behaviour that might not otherwise exist (Langan-Fox et al., 2000). This criticism is founded in seminal research by Nisbett & Wilson (1977) which demonstrated that individuals have little direct access to higher order cognitive processes. However, verbal reporting should not be confused with introspection. Chi (2006) describes the former as problem centred and outward looking, and the latter as concerned with one's own thoughts and experiences. There is considerable evidence for the validity of protocol analysis when the question of interest is restricted to the representation of information, and when appropriate elicitation techniques are used (Ericsson & Simon, 2003).

3.4.4 Conceptual Methods

Conceptual methods (or more accurately scaling methods) produce representations of domain concepts and their structure or interrelations (Cooke, 1994). The general steps required are: (1) elicitation of concepts through interview or analysis of documentation, (2) collection of relatedness judgments from experts, and (3) reduction and representation of relatedness data

(Cooke, 1999). Common methods of data reduction and representation include cluster analysis, multidimensional scaling, and the Pathfinder technique. Broadly speaking, these methods use pairwise estimates to generate a spatial or graphical representation of the concepts, represented by nodes, and their relations, represented by weighted links between nodes (Cooke et al., 2000).

The utility of methods such as Pathfinder lie in the fact the representations can be compared both qualitatively and quantitatively across groups and individuals. In Pathfinder, for example, an estimate of mental model similarity can be obtained by dividing the number of common links in two networks by the number of links that are in either of the networks (Langan-Fox et al., 2000). Applied examples of the Pathfinder technique are provided in Cooke et al. (2000) and Langan-Fox et al. (2000).

There are several disadvantages to conceptual methods of mental model elicitation and representation. First, the concepts used for relatedness judgments are typically experimenter-generated, which distances the analysis from the participant's actual representation and detracts from the ecological validity of the analysis. Second, these methods are indirect in that knowledge representation is inferred through judgments of conceptual relatedness (Cooke, 1999), with the conceptual map representing associative but not semantic information – essentially the closeness of concepts to one another (Gordon et al., 1993). The obtained structure does not represent knowledge in a causal or hierarchical manner – it is in essence a function of the elicitation technique, and may have little direct relation to task performance (Langan-Fox et al., 2000). As a consequence, the results of these analyses do not lend themselves to the development of training programs (Gordon et al., 1993). However, Cooke et al. (2004) argue that the representations produced by indirect methods may be more effective predictors of team performance than those produced by direct methods precisely because of their indirect nature: individuals with “shared” mental models may not know the identical facts and rules in relation to a task, but they may have the same abstract or higher order understanding of a task (as reflected in relatedness judgments).

3.4.5 Issues in Eliciting and Measuring Knowledge

In selecting an elicitation method, Hoffman and Lintern (2006) suggest the consideration of differential utility: each knowledge elicitation method has its strengths and weaknesses, and any given method might be more useful for certain purposes or applicable to certain domains. For example, the assessment of rapidly changing situation models is most appropriately conducted using process-tracing approaches. Typically, more than one elicitation method is used over the course of a single study.

It is important to keep in mind that, in relation to team mental models, it is neither necessary nor feasible to measure everything that each team member knows. Cooke et al. (2000) characterise the target of measurement, in a given task and environment, as the knowledge required by interdependent team members and the distributional characteristics of that knowledge across the team.

It is also important to take into account the fact that team knowledge does not exist within a static environment: team situation models are constantly changing, and team mental models, although more stable, will also evolve over the life of the team. As the expertise of team

members increase over time, their mental models should in fact become more abstract, more hierarchical in nature, and more coherent and complete (Chi, 2006). Consideration should be given, therefore, to the point at which team knowledge is elicited. The development of team mental models can be assessed by obtaining measures at appropriate discrete points during team performance. For some forms of knowledge, such as situation models, it may be possible to collect data continuously and synchronously with task performance using video records, event logs, and communication records (Cooke et al., 2000).

3.5 Team Knowledge Metrics

In order to assess the similarity of knowledge within a team, to aggregate individual results to generate a representation of team knowledge, or to assess the accuracy of individual or team knowledge it is necessary to quantify the information obtained through the knowledge elicitation procedures described above (Cooke et al., 2000). The first step is the development of a coding scheme (see Section 3.4.3). Cooke et al. describe a variety of metrics that can then be employed in evaluating team knowledge: similarity, accuracy, interpositional accuracy, and knowledge distribution metrics.

Similarity metrics have been the focus of most research, and have been used to investigate the existence of 'shared' mental models from an actuarial perspective. Depending on the nature of the obtained data (and of course the heterogeneity of team knowledge), similarity can be measured in terms of the number or percentage of responses (or concepts, and relations between concepts) that are identical for team members. The literature provides little direction as to the degree of overlap among individual mental models that needs to exist before they are considered shared, or the proportion of members of a team who must be congruent in their thinking for a shared mental model to exist (Klimoski & Mohammed, 1994).

Accuracy metrics provide more informative data than similarity metrics. As Cooke et al. (2000) point out; it is possible for members of a homogenous team to possess identical, but completely inaccurate knowledge. Conversely, members of heterogeneous team could possess perfectly accurate but completely dissimilar knowledge. A simple measure of team accuracy is the average number or percentage of valid responses or concepts. Generally, the elicited knowledge (such as a conceptual graph or a Pathfinder network) is compared with a referent obtained, by theoretical or empirical means, from domain experts. For heterogeneous teams, Cooke et al. recommend partitioning the knowledge base into chunks associated with each member's team role before calculating accuracy metrics.

Closely related to heterogeneous accuracy is the concept of interpositional accuracy, which concerns the knowledge associated with the roles of other team members. Interpositional accuracy can be obtained by comparing a team member's knowledge with the referent of the other team members (Cooke et al., 2000).

Cooke et al. (2000) point out that accuracy metrics, although providing a measure of the extent to which individual team members have mastered portions of the knowledge base, do not reveal gaps in the team knowledge. They recommend using knowledge distribution metrics, which refer to the manner in which specific knowledge is distributed among members of a team. To calculate these metrics, it is necessary to establish a set of items (or concepts) which completely cover the domain in question. It is then possible to calculate the percentage of the

knowledge base covered by at least one team member (and also the percentage of knowledge base that is redundantly distributed).

When combining metrics across team members in order to produce a metric that is representative of the team, it is important to consider the most suitable aggregation method (Cooke et al., 2000). This will depend on the nature of the data; various options include the mean, median, sum, minimum or maximum (which will produce values that are a function of the weakest or strongest team member), or the responses given by a certain proportion of team members (similar to the prototype approach, e.g., Rosch & Mervis, 1975).

4. Measuring Cognitive Factors: Situation Awareness

In complex and dynamic environments – such as aviation, air traffic control, and military command and control units – operators must attend to, and maintain awareness of, many different aspects of their environment in order to perform effectively. The mental model that represents an individual's understanding of the current situation and its implications is commonly referred to as their *situation awareness* (SA; Vidulich, 1995). SA, referred to in the previous section of this report as a *situation model*, is widely touted as a critical foundation for sound decision making in environments characterised by multiple, complex, and constantly changing sources of information (Endsley, Selcon, Hardiman, & Croft, 1998; Vidulich, 1995).

The most widely-cited conceptualisation of SA is Endsley's three stage model (1990), in which SA is defined as the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future. As this definition suggests, Endsley's model separates SA into three distinct levels. Level 1 SA involves attending to, and perceiving, elements of the environment deemed important according to the goals and objectives held in working memory. Newly apprehended information, held in working memory, must then be integrated with existing knowledge, in the form of schemata or mental models, to comprehend the developing situation (Level 2 SA). This understanding can then be used to forecast possible future events (Level 3 SA). According to this model, the processing of information in order to achieve a higher level of SA results in an increase in the quality or richness of one's awareness (Stanton, Chambers, & Piggott, 2001).

4.1 Team Situation Awareness

In the applied literature, SA is usually considered at the level of the individual. The SA of individual operators, however, is not sufficient to characterise the SA of a team. Invariably, individuals must share their situation awareness appropriately with other team members in order for the team as a whole to perform successfully (Vidulich, Bolia, & Nelson, 2004). In keeping with the patterns evident in the broader team knowledge literature, there is considerable diversity among conceptualisations of the team SA construct (Sarter & Woods, 1991; Flach, 1995; Pew, 1995), and the same or similar constructs have appeared in a variety of guises (see Jentsch, Barnett, Bowers, & Salas, 1999; Krauss & Fussell, 1990; Orasanu, 1994, 1995; Regal, Rogers and Boucek, 1988; Roberston & Endsley, 1995; Sarter & Woods, 1997; Wellens, 1993). In this regard, much of the discussion in Section 3.2, regarding the concept of team knowledge, applies to the concept of team SA.

One commonality to definitions of team SA is a reference to individual SA (Salas et al., 1995). When working in a team, each team member has their own roles and goals, which contribute to the overall team goals. Associated with each team member's roles and goals is a set of SA elements – aspects of the task environment about which that member is concerned. Endsley (1995a) has described team SA as the degree to which every team member possesses the SA required for his or her responsibilities. Endsley stressed that team SA is independent of any overlap in the SA requirements of team members: if two team members (for example a pilot and co-pilot) both need to know certain information, it is not sufficient that one (the co-pilot) possesses this information and the other (the pilot) does not. The advantage of Endsley's (1995a) definition of team SA – which equates to the sum of each individual team member's SA – is that, given a valid measure of individual SA, team SA is also amenable to measurement (Bolia & Nelson, 2007). Bolia and Nelson suggest, however, that there are possibilities for the expression of team SA other than a simple additive function. One option would be to consider team SA as the sum of the SA for each individual team member, with individual's SA scores weighted to reflect their contribution to overall mission success. Another option would be to use a weighted multiplicative rather than an additive function of individually measured SA.

Prince, Ellis, Brannick, and Salas (2007) agree with Endsley's (1995a) definition of team SA, but suggest that it be amended to include an explicit reference to the need for team members to also possess the situation awareness required for the team's responsibilities. This point has much in common with the notion of *shared* SA.

Like team knowledge, team SA can be homogenous (shared) or heterogeneous (distributed) in nature (Bolia & Nelson, 2007). In most circumstances, members of a team will have some similar SA requirements, with shared SA an important factor when consistency between the SA of individual operators is necessary for achievement of the team goal. Consistent with the notion of a shared mental model, shared SA is conceptualised as the shared understanding of a situation among team members at one point in time⁷ (Salas et al., 1995). Kaber and Endsley (1998) state that shared SA can occur in various states within a team: (1) all team members may have the correct assessment of the situation, (2) all team members may have an incorrect assessment of the situation, or (3) all team members may have a different assessment of the situation, some with correct and some with incorrect assessments of the situation.

Fundamental to shared SA is another commonality to definitions of team SA: a reference to team processes, in particular communication (or information sharing) and coordination (Salas et al., 1995). According to Salas et al. (1995), just as specific cognitive processes (i.e., perception, comprehension, and projection) underlie Endsley's (1995a) model of individual SA, specific team processes underlie shared (and therefore team) SA. Kaber and Endsley (1998) in fact suggest that team coordination primarily consists of the sharing of SA on elements of mutual interest, and Endsley (1995a) is of the view that the degree to which each team member has accurate SA on shared items could serve as an index of the quality of team communication.

⁷ Salas et al. (1995) used this description to describe team SA.

The pure conceptualisation of shared SA is the intersection of each individual's SA. Bolia and Nelson (2007) point out that if this strict definition of shared SA is used to operationalise team SA, then team SA will be zero if just one team member lacks SA. Bolia and Nelson (2007) argue that this is inconsistent with the notion that a large team may still have good SA, even if one or more of the team members have none; they propose that team SA is best viewed as existing on a continuum between distributed and shared.

It is important to keep in mind that the nature of the particular team under consideration will determine the conceptualisation and operationalisation of that team's SA (and shared SA) needs. Factors such as team member roles, the degree of interdependence within the team, the unique task environment, and the specific task goals will dictate the SA requirements of each team member, which will in turn have implications for team SA requirements (Kaber & Endsley, 1998; Salas et al., 1995).

4.2 The Measurement of Individual Situation Awareness

There have been a number of different approaches to the measurement of an individual's SA, each with their own strengths and weaknesses. Broadly speaking, these SA metrics can be characterised in terms of whether they adopt an indirect or a direct measurement approach. Accordingly, the following discussion considers these two approaches separately.

4.2.1 Indirect Measures of Situation Awareness

Situation awareness is a psychological construct, and not directly observable. Consequently, many attempts to measure SA have employed an indirect approach. These indirect (or implicit) approaches involve recording some readily observable variable which is expected to correlate robustly with SA – typically, a measure of task performance. Recently, attempts have also been made to identify physiological correlates of SA. These approaches will be discussed in turn.

4.2.1.1 Performance Measures

The basic assumption underlying performance measures of SA is that the level of SA of the operator will be reflected in his or her performance on the task (Breton & Rousseau, 2003). Performance measures can be either global or specific in nature. In a flight simulation task, for example, a global measure of SA would be drawn from variables related to overall mission performance, whereas a specific behavioural measure of SA would be drawn from the pilot's ability to maintain a desired altitude, or respond in a timely and appropriate manner to a particular warning. By their nature, performance measures are concerned with measuring the end-product of SA rather than the three individual levels of SA outlined in Endsley's (1990) model.

An example of the use of indirect performance measures can be found in Gugerty's (1997) investigation of people's ability to monitor changing spatial information in a simulated driving task. The task required participants to maintain speed and direction in traffic unless they needed to avoid a hazard, in which case they were permitted to accelerate, brake, or change lanes. During each trial a vehicle would pull into the participant's lane, forcing the participant to take evasive action in order to avoid a collision. However, during catch trials, a car would be located in the driver's blind spot. The indirect performance measures of SA that

were used included (1) whether or not the driver changed lanes into the blind-spot car, and (2) the response interval between presentation of the initial hazard and the driver's response.

Sarter and Woods (1991) developed a variant of indirect performance measurement known as external task performance measurement. In this method a piece of information is removed, changed, or added during a task without the operator's knowledge; SA is operationalised as the time it takes for the operator to react to this change. Sarter and Woods (1997) used this method in a flight simulator study investigating pilot awareness of automation activities. The study involved a number of civil aviation scenarios in which pilots were required to follow regular approach and landing procedures at a busy airport. One scenario, for example, presented a loss of a navigation signal on approach – a reason for immediate go-around. The quality of the participant's SA was inferred from the time between the loss of navigation signal and the pilot's initiation of a go-around.

The key benefits of performance measures of SA are that they are objective, non-intrusive, and relatively easy to employ. However, performance measures provide little insight into the processes used to build and maintain SA, or achieve the task goal (Endsley, 1995b). There is also an issue of circularity of reasoning, in that we 'know' an operator possessed a good level of SA because they performed well, and they performed well because they possessed a good level of SA (Flach, 1995). A second and related criticism stems from the fact that many different factors influence performance on a task. That is, poor performance may not necessarily be the result of poor SA (and vice versa).

4.2.1.2 *Physiological Measures*

Psychophysical methods such as eye-tracking and electroencephalography, which attempt to assess physiological correlates of cognition, offer potential as a means of quantifying SA. Eye-tracking technology measures the direction and duration of a participant's gaze and has been used to assess SA in both the aviation (see Alexander & Wickens, 2005; Wickens et al., 2005) and driving (see Harbluk & Noy, 2002) domains. The rationale behind this technique is that the content of an individual's gaze is related to their awareness of the environment. Eye-tracking data can be collected continuously during task execution, making it an attractive prospect for SA research. However, the direction of a person's gaze does not necessarily equate to a direction of attentional resources, and it is therefore difficult to infer a one-to-one relationship between gaze and SA. Furthermore, shared perception does not equate to shared interpretation (Bolia & Nelson, 2007). Salas, Prince, Baker, and Shrestha (1995) highlight the importance of drawing a distinction between situation assessment and situation awareness: the former is a process, and the latter a state. Situation assessment results in knowledge that *may* become part of situation awareness (Sarter & Woods, 1991). Electroencephalography (EEG) is less intrusive than many eye-tracking techniques, and also presents the opportunity to collect continuous data. Evidence for the validity of this technique was provided by French, Clarke, Pomeroy, Seymour, & Clark (2003), who found correlations between EEG response patterns and subject matter expert ratings of the SA of participants playing a PC-based war game. While the method used by French et al. shows the promise of EEG techniques for the assessment of SA, this is an area in which further research is required.

4.2.2 Direct Measures of Situation Awareness

Direct (or explicit) measures of SA are designed to address the subject's understanding of the situation and its implications in a transparent and unmediated way. Direct measures can be divided into two sub-categories: subjective and objective.

4.2.2.1 Subjective Measures

Subjective measures of SA require operators to rate their personal experience of SA along various dimensions either during, or shortly following, task performance. As subjective measures of SA are usually gathered via survey methods, they have the advantage of being inexpensive and easy to administer.

One of the most widely used subjective measures of SA is the Situational Awareness Rating Technique (SART; Taylor, 1990). There are two forms of the SART: the ten dimensional (10-D) scale and the three-dimensional (3-D) scale. The 10-D scale contains one item relating to each of ten independent bipolar dimensions derived, using principle components and cluster analysis, from an initial set of 44 constructs identified as relevant to SA by test aircrew. These ten dimensions are arranged, as illustrated in Table 4.1, in three generic groupings: (1) demand on attentional resources, (2) supply of attentional resources, and (3) understanding of the situation. The system operator indicates rates each of the items on a seven point scale (1=low, 7=high).

Table 4.1. Generic grouping of items in the 10-D SART scale

Demand on attentional resources	Supply of attentional resources	Understanding of the situation
Instability of situation	Arousal	Information quality
Variability of situation	Spare mental capacity	Information quantity
Complexity of situation	Concentration	Familiarity
	Division of attention	

Note. Adapted from "Evaluation of the Situational Awareness Rating Technique (SART) as a tool for aircrew systems design, by S. J. Telcon and R. M. Taylor, 1990, in *Situational Awareness in Aerospace Operations*, p. 5-2.

The 3-D scale was developed for use in testing situations where there may be constraints on participants' time (such as on a flight deck). In this variant of the SART, participants are asked to respond to a single item representing each of the three generic grouping dimensions described above, as well as a fourth item – "situation awareness" – that represents the broader construct. This version employs a visual analogue scale, with participants asked to rate each of the dimensions by placing a mark on a continuous, 100 mm line between the verbal anchors "Low" and "High". Flight simulator studies have provided evidence of the construct validity of the SART, with ratings correlating with measures of operator performance (faster response times to warning signals; Selcon & Taylor, 1990) as well as other subjective measures of situation awareness (Endsley, Selcon, Hardiman, & Croft, 1998).

The major strengths of the SART are that it is easy to administer and score, and is applicable to a wide variety of environments – including real world tasks and simulations (Endsley et al.,

1998). However, a number of criticisms have also been levelled at this assessment technique. First, the SART has been criticised on the basis that operators are poorly placed to assess their own SA, due to their inability to 'know what they don't know' (Endsley, 1995a). As pointed out by Vidulich (1995), a pilot may rate his or her SA as high immediately prior to being surprised and shot down by an unobserved enemy. A second problem with self-ratings such as SART is that it is impossible to determine whether the high correlations that have been observed between self-rated SA and task performance (e.g., Venturino, Hamilton, & Dvorchak, 1990) arise because SA is a good predictor of task outcomes, or alternatively, because participants have difficulty distinguishing between evaluations of their own SA and those of their own performance (Endsley, 1995b). A third criticism of SART is that this technique may confound assessment of SA with workload issues (the supply and demand of attention). In particular, Endsley (1993) argues that participants may have difficulty distinguishing between SA and workload, even though SA may in fact be independent from workload in many situations.

4.2.2.2 *Objective Measures*

Objective measures of SA are designed to be independent of the subjective impressions of both the participant and the administrator. One approach that meets this criterion is the use of probing questions to assess an operator's awareness of particular items within the system, the responses to which are then compared with records of the true system state. A significant strength of the memory probe technique is that it does not rely on subjective assessments of SA, or on an interpretation of complex behaviour (Endsley, 1995b). Problems arise, rather, with the practicalities of implementing this technique. First, retrospective (post-test) implementation places a heavy reliance on the memory of the operator, which may confound the measurement of SA. In particular, operators tend to forget their misperceptions over the course of the task, and are therefore only able to provide insight into their understanding of the event as a whole, rather than the component elements of their SA (Endsley, 1995b). An alternative to post-test implementation is on-line implementation, in which operators are required to respond to questions while they are actively involved in the task. A drawback to this technique is that responding to the questions is likely to add to operator workload, and may in itself have a negative impact on SA. A third technique is implementation during freezes. Although this technique overcomes some of the problems associated with retrospective and online techniques, an obvious disadvantage is that freezes are usually only possible in simulated tasks, which limits the ecological validity and generalisability of research conducted in this manner. A further drawback is that freezes are extremely intrusive, serving as a significant source of annoyance to participants (Brickman et al., 1995).

The most widely used objective technique for measuring SA is the Situation Awareness Global Assessment Technique (SAGAT; Endsley et al., 1998). The SAGAT is an objective technique that utilises periodic, randomly-timed freezes in a simulation scenario, during which the operator's displays and window views are momentarily blanked. The operator is then required to respond to a set of queries regarding his or her knowledge of what was happening at the time of the freeze (Endsley et al., 1998). Responses to queries are scored with regard to the true state of affairs at the time of the freezes. The queries typically cover the three levels of SA described in Endsley's (1990) model: perception, comprehension, and projection. Unlike the fixed items used in SART, SAGAT queries must be tailored to each domain using an in-depth cognitive task analysis. Thus, although the SAGAT is broadly applicable, it requires considerable customization for each specific use. A methodological drawback to the SAGAT is

that it generates binomial data (responses are either correct or incorrect), which necessitates the use of statistical techniques that are low in power. The SAGAT has been used in domains such as aircraft cockpits (Endsley et al., 1998), air traffic control (Endsley & Rodgers, 1996), maintenance systems (Truitt & Ahlstrom, 2000) and nuclear power plant control (Hogg, Torralba & Volden, 1993).

A good illustration of the use of the SAGAT is Endsley et al.'s (1998) investigation of the effect on fighter pilot SA of a system for directly presenting information on threat aircraft capabilities. This system was designed to provide pilots with the information required for replanning around pop-up threats. Pilots completed a part-task mission simulation which involved following way-points in a low-level ingress flight. During the flight pop-up threats appeared, which the pilots were required to avoid while staying as close to their pre-determined flight path as possible. In the control condition, only the threat aircraft symbol and supporting information (aircraft type, speed) were shown. In the experimental condition, the same information was supplemented with a graphical display of the threat aircraft's launch success zone. Performance was measured in terms of deviations from the assigned course, and total time spent inside the threats' launch engagement zones. The SAGAT was administered during freezes which occurred at random times during the task. Endsley et al. were able to use the SAGAT to demonstrate the effects of the launch engagement zone display at each of the three levels of Endsley's (1990) model of SA. For example, the experimental condition was associated with reduced level 1 SA regarding the location of threat aircraft, but increased level 1 SA regarding own-aircraft heading and roll attitude. In terms of level 2 SA, participants had a better awareness of the imminence of threats. In terms of level 3 SA, participants demonstrated a reduced awareness of the most likely threats on their current course.

This study is significant in that it also involved a comparative evaluation of the SART and the SAGAT. The most notable finding in this respect was that the SART and SAGAT measures were not correlated, with each providing qualitatively different information about the effect of the experimental manipulation on pilot SA. Endsley et al. (1998) concluded that the SART may be most useful in terms of providing information on the *perceived* quality of SA, which may be instrumental in determining how an operator will act on his or her SA in a given situation. The SAGAT, however, was seen as a more useful diagnostic instrument, and a more veridical representation of SA.

4.2.3 The Global Implicit Measure

In order to assess pilot SA in a simulated air combat task, Brickman et al. (1995) employed a technique intended to incorporate the strengths of implicit performance-based approaches whilst minimising the weaknesses of explicit memory probe approaches. This Global Implicit Measure (GIM) approach compares measures of human performance to previously defined behavioural constraints, referred to by Brickman et al. (1995) as "rules of engagement" (ROE). These ROE are developed through a task analysis conducted in association with a subject matter expert. This task analysis is a critical aspect of the GIM approach, which rests on the assumption that, in the case of complex and dynamic task environments, the relevance of any particular stimulus to an operator's situation awareness is context dependent – and will therefore vary as a function of the specific task, the environment, and the tactical objective (Sarter & Woods, 1991). In the air-to-air combat scenarios employed by Brickman et al. (1995),

the ROE (which are expressly communicated to operators) included parameters such as speed, altitude, course, radar mode, and weapons selection. In the GIM approach functionally adequate SA is argued to occur when objectively measured aspects of human performance are in accordance with the previously identified parameters. Measures of operator-system performance are continuously recorded on-line, and each implicit measure is scored at a rate equal to the frame rate of the simulation: a score of “1” indicates adherence to the ROE, and a score of “0” indicates a failure to adhere to the specified ROE. This enables the calculation of a proportion score reflective of the operator’s SA in relation to a particular task segment or goal, which can then be combined with other data from similar implicit probes to obtain measures of SA for the overall task (Vidulich, 1995). Brickman et al. (1995) and Vidulich (1995) cite a number of benefits of the GIM approach: (1) the detailed task analysis lends high operational (or ecological) validity, (2) it is non-intrusive, (3) the focus on objective and quantifiable measures of human performance allows for precise measurement, and (4) the scoring method provides sufficient data to run statistical analyses possessing reasonable power.

4.3 The Measurement of Team Situation Awareness

Over a decade ago, Salas et al. (1995) described the measurement approaches designed to assess individual SA as inadequate for assessing team SA, and made the point that the lack of well-developed team SA measurement tools was hindering the development of instructional strategies for team SA (see also Brannick, Prince, Prince, & Salas, 1995). More recently, Prince, Ellis, Brannick, and Salas (2007) have reiterated this point, stating that effective training for team SA cannot be implemented in the absence of a valid, reliable method of measuring team SA – such a method is necessary for initial assessment of team SA, evaluation of training benefits, and feedback to team members.

Salas et al. (1995) and Vidulich et al. (2004) argue that, at the very least, the assessment of SA in teams requires an assessment of each individual’s SA, as well as an assessment of how that SA is shared within the team (through team processes). Vidulich et al. (2004) argue that the most efficient way to study SA processes in a team environment is to investigate the pattern and content of communications between team members during mission performance (see Section 2.6, above).

To date, however, most of the research investigating team SA has employed largely the same techniques as those discussed in the context of individual measurement. This data is simply treated differently, with team SA usually operationalised as either the average SA of all individual team members (e.g., Sebok, 2000) or the sum of the SA held by individual team members (e.g., Prince et al., 2007). This simple aggregation of individual SA is not consistent with the notion that team SA is more than the sum of the SA of individual team members (Bolia & Nelson, 2007; Endsley, 1995a; Salas et al., 1995), and it is particularly inappropriate for heterogeneous teams (see the discussion in Section 3.2, above). Although team SA must arise from some combination of the SA possessed by the individual team members, it is a more complex, and perhaps qualitatively distinct, construct.

The research described below is representative of current approaches to the assessment of team SA. Note, however, that much of the previous discussion on the measurement of team knowledge (Section 3.4) and team knowledge metrics (Section 3.5) has relevance for the measurement of team SA.

4.3.1 Examples of Aggregation Measures of Team Situation Awareness

Sebok (2000) employed direct measurement techniques with an 'average' conceptualisation of team SA in a simulation study of the effects of interface design and staffing levels on team SA in a nuclear power plant. Teams of operators participated in complex scenarios, working in either a simulated conventional plant with a hard-control interface, or in a simulated advanced plant with a computerised interface. The computerised interface was designed to provide integrated information and a common overview display, which was expected to support team SA. Two team staffing levels (normal and minimum) were assessed in each plant condition. Team SA was assessed with the SAGAT modified for the process control industry. The SA questions asked of participants were those in the Situation Awareness Control Room Inventory (SACRI) developed by Hogg, Folleso, Torralba, & Volden (1995). These questions are specific to nuclear power plant control and relate to the trends of different process parameters. Responses to these questions were compared against simulator records to verify correct and incorrect responses, and the scores were then averaged across team members to obtain a measure of team SA. Through this technique, Sebok (2000) was able to determine that in the conventional plant condition the normal sized team had a higher average level of team SA, while in the advanced plant condition the reduced member team had a higher average level of Team SA.

An example of the additive conceptualisation of team SA is Prince et al.'s (2007) examination of SA in two-person cockpit teams, which purported to compare indirect (which they label implicit) and direct (which they label explicit) techniques for measuring team SA. In the first phase of the experiment, teams flew a complex mission in a high-fidelity, full-motion simulator. Two experienced military instructors provided ratings of performance and situation awareness. Team performance was evaluated in relation to problem oriented criteria, which consisted of crew member's handling of icing, boost pump, and electrical fire problems according to standard procedures. Team situation awareness was evaluated using what was described as a process oriented indirect approach. Although no further information was given, this measure appears to have consisted simply of a subject matter expert rating, on a 5-point scale, of crew member's SA (and could therefore reasonably be considered as an assessor rating of explicit team SA). In the second phase of the experiment, which took place 15 minutes after debriefing of the first mission had ended, the same teams flew a simple mission in a low fidelity simulation environment. SA was evaluated using freezes and probe queries, consistent with the SAGAT method (direct measurement). Each team member received a score according to how many items they answered correctly. Team SA was operationalised as the sum of these two scores. The authors acknowledged that there may be other approaches to aggregating individual SA, conceding that the implications of one, both, or neither crew member having particular flight knowledge may be quite different. Performance measures were not obtained in the low-fidelity scenario.

Prince et al. (2007) found that, in the high-fidelity scenario, the indirect measure of SA correlated strongly with the problem-oriented performance measures: the smallest value of Pearson's r was .62. The low-fidelity scenario direct measures of team SA showed a weak and inconsistent relation with the high-fidelity scenario performance measures: values of Pearson's r ranged from .13 to .53. The relation between the indirect and direct measures of team SA was also modest, but significant, with $r = .43$. The unconvincing relations between direct team SA measures and performance, and between the direct and indirect measures of

team SA, is not surprising. An obvious (and significant) problem with this study is that scores used to compare indirect and direct methods of measuring SA were collected during the execution of separate missions. The correlations between measures of SA therefore confound differences in measurement method with measurement occasions (phase one versus phase two), mission scenario (point-to-point flying versus problem management), and work environment (high versus low fidelity simulation). In addition, the relatively stronger relationship between assessor-rated team SA and assessor-rated performance scores is suggestive of method effects (i.e., an artificial inflation of the apparent relationship between two variables due to commonalities in the method of their collection).

Problems in experimental design aside, the weak and inconsistent relations between direct and indirect measures of SA in Prince et al.'s (2007) study are understandable in terms of the differences in what is actually measured using these two techniques. In particular, whilst a good understanding of the elements of the environment may be one determinant of the actions upon which indirect measures are based, it is not the only determinant. As noted earlier, indirect measures of SA inherently confound assessment of SA with assessment of a myriad of other factors that can affect the actions that are taken in a particular circumstance (e.g., workload). This applies to measurement of SA at the team level just as it applies at the individual level.

4.3.2 Measures Specific to Team Situation Awareness

Gorman and colleagues (Gorman, Cooke, & Winner, 2006; Gorman, Cooke, Pederson, Connor, & DeJoode, 2005) have developed an approach to measuring team SA called coordinated awareness of situations by teams (CAST), intended primarily for use in distributed command and control (network-centric) environments – which typically involve large teams comprised of members with specialised roles. Gorman et al. (2005) posit that in this type of environment the view of team SA as a shared understanding between team members is neither feasible nor desirable; it may in fact be prohibitive and counteractive to provide all team members access to the same information. As team members are not expected to have a complete tactical picture of the environment at all times, considerable interaction between team members is required in order to accomplish the team's goals (Gorman et al., 2006).

The CAST approach is inspired by ecological views on psychological phenomena, and focuses on the interaction processes in teams. Gorman et al. (2006) make the assumption that team environments are inherently dynamic, and that team coordination is constantly changing in pace with the environment and in line with the team's goals. They argue that it is not enough to assess team SA during this typical task performance. Rather, the optimal context in which to assess team SA is when the typical dynamic environment itself undergoes change, forcing, in their words, a "change in changing coordination" (p. 1315) in order for the team to maintain progress towards the achievement of their goals. The changes in the dynamic team environment that necessitate a change in team coordination processes are referred to as 'roadblocks,' an example of which would be an unanticipated disruption to normal communication channels. In this paradigm, each team member has their own local perception of the roadblock. However, a coordinated perception is required in order to fully comprehend and act upon the roadblock (Gorman et al., 2005).

Table 4.2. Central concepts of the coordinated awareness of situations by teams (CAST) approach to measurement of team SA

Identify roadblocks	Unlikely or unanticipated events requiring adaptive and team-level solutions.
Document primary perceptions	One or more team members perceive some aspect of the roadblock and can react to it.
Document secondary perceptions	An operator attuned primarily to his or her own perspective on the roadblock is exposed to a different perspective on the roadblock through interactions with operators more attuned to other aspects of the environment.
Document coordinated perceptions	Involves the reciprocal effort involved in team members explaining what they are perceiving to each other in order to create a picture of the situation that is more than just the sum of their individual awareness
Document coordinated actions	Involves the coordinated actions necessary to overcome specific roadblocks, within the constraints imposed by the situation and other team member activities.

Note. Adapted from "Measuring team situation awareness in decentralized command and control environments," by J. C. Gorman, N. Cooke, and J. L. Winner, 2006, *Ergonomics*, 49, p. 1317.

Measurement takes place when operators encounter novel or unlikely situational constraints (roadblocks), and involves documenting the actions, comments, behaviours, and interactions of team members in the context of the specific task environment and team goals (Gorman et al., 2006). Team SA is measured as the team's coordinated response to situational change. The five central concepts that form the basis of a CAST methodology and instrument are presented in Table 4.2.

Gorman et al. (2005) utilised the CAST approach to compare the team SA of co-located and distributed 3-person teams operating in an unmanned air vehicle (UAV) synthetic task environment. Each team member had a specialised role (pilot, photographer, or navigator) and coordinated their activities over microphones and headsets in order to fly their UAV to take photos of targets. Mission performance scores were based on a weighted sum of variables including number of photos, mission time, and time spent in an alarm state. The roadblock was a 5-minute communication channel glitch from the navigator to the pilot (pilot could not hear navigator, but the navigator could hear the pilot). The CAST scoring procedure consisted of monitoring communications around this roadblock and scoring performance on four dimensions: (1) which team members independently noticed the glitch, (2) which team members discussed the glitch, (3), actions taken in order to circumvent the glitch, and (4) whether or not team members overcame the roadblock. The first three dimensions relate to the documentation of primary perceptions, coordinated perception, and coordinated actions described in Table 4.1. Overall, distributed teams were found to perform better than co-

located teams. Of more interest, Gorman et al. (2005) found that high performance was predicted by higher levels of first perception of the glitch, along with lower team member involvement in coordinated perception and action. Gorman et al. interpreted this finding as contrary to the view that team SA is optimal when all members share a common picture, and supportive of the argument that a shared situation model can underlie maladaptive team behaviour (in this case, by unnecessarily increasing workload).

The CAST approach has much in common with the assessment of communication behaviours discussed earlier, in Section 2.6. An example of communication-based assessment of team behaviours, framed explicitly in the context of SA, is Burke, Murphy, Covert, and Riddle's (2004) investigation of the situation awareness in teams conducting an urban search and rescue disaster response drill with teleoperated robots. In this study, a content analysis of the interactions (verbal communications) between team members provided the principal source of data for measures of SA. Each statement was coded on four categories: (1) speaker-recipient dyad, (2) form or grammatical structure of the communication, (3) function or intent of the communication, and (4) content or topic of the communication. This coding scheme enabled Burke et al. to determine which team members were interacting, what they were communicating about, and provided indicators of operator situation awareness at the three levels of Endsley's (1990) model. The most relevant finding was that team processes and communication were critical to improving situation awareness in circumstance in which it was degraded.

Gorman et al. (2006) argue that memory-probe techniques such as the SAGAT, while appropriate for measuring knowledge-based constructs, are not appropriate for measuring highly dynamic phenomena such as SA – particularly if the real goal is to measure the cognitive processes *underlying* the ongoing awareness of an operator. In this sense, they consider the so-called 'direct' SA measures such as the SAGAT to be an indirect form of measurement. Gorman et al. argue that the CAST measurement approach is more direct, and less inferential, than knowledge-based elicitation approaches (e.g., the SAGAT).

Gorman et al. (2006) cite a number of advantages to the CAST approach. First, unlike memory probe approaches to the measurement of SA, which are not only intrusive and interfere with ongoing team SA processes but are also both retrospective and introspective, CAST measures can be obtained whilst the task is in progress and without disrupting normal team processes. Second, team interactions are more readily observable than the individual or aggregate knowledge (situation models) of team members, and as a consequence they provide the opportunity of observing team cognition directly. Finally, Gorman et al. criticise memory-probe techniques on the grounds that they are outcome oriented, and that the results of assessments do little to assist in the diagnosis of the processes underlying poor team SA. They describe the CAST approach, on the other hand, as process oriented and thereby capable of providing information about the process by which the goal of a common picture is achieved.

The CAST philosophy also has implications for the training of team SA; in particular, it calls for training that is focused on the coordinated interactions between team members (Gorman et al., 2006). Synthetic training environments, which allow for the control of trigger events (i.e., roadblocks) used to provide observation points, appear particularly suited in this respect (see also the discussion on the TARGETS/EBAT approach in Section 2.5). Gorman et al. also conclude that their research indicates that team SA may be improved through the use of

tools that facilitate the adaptive and timely exchange of information among team members, so that the right information is communicated to the right person at the right time.

4.3.3 Team Mutual Awareness

Team Mutual Awareness (TMA; MacMillan et al., 2004) refers to the extent to which individual team members are aware of other team members' behaviours. The concept of TMA is similar to team SA in that it concerns dynamic knowledge that develops and evolves during the course of the task. It differs, however, in that it refers explicitly to states and processes *within* the team, rather than to elements of the environment (which are usually conceived of as external to the team). MacMillan et al. put forward a model of TMA and described techniques used to measure its components. According to MacMillan et al., in order for teams to operate effectively, team members must be aware of three factors underlying the actions of their team mates (see the model of TMA in Figure 4.1).

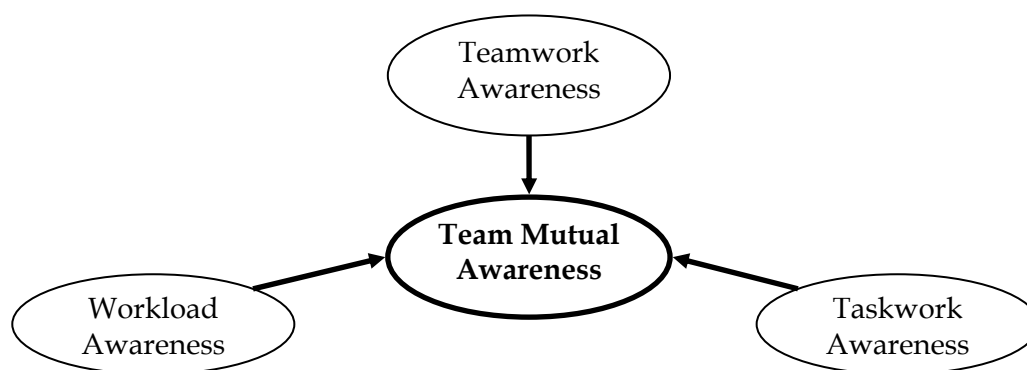


Figure 4.1 The components of Team Mutual Awareness (TMA). From "Questionnaires for distributed assessment of team mutual awareness," by J. MacMillan, M. J. Paley, E. B. Entin, and E. E. Entin, 2004, in N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.) *The Handbook of Human Factors and Ergonomic Methods*, p. 51-1. Boca Raton, FL: CRC Press.

In this model, Taskwork Awareness refers to an awareness of what tasks other team members are completing. Workload Awareness refers to an awareness of the cognitive load imposed on team members by the taskwork they are undertaking. Teamwork Awareness refers to an awareness of the execution of teamwork behaviours, such as those described in Section 2 of this report. MacMillan et al. outlined questionnaires that they had used in measuring each of the components of TMA. These questionnaires purportedly tap into mutual awareness by asking team members to indicate not only their experiences of aspects of the team's performance, but also their perceptions of the experiences of others. Using this method it is possible to compare, for example, a team member's reported workload with the workload that others in the team perceived that he/she was experiencing. A high level of understanding in relation to these factors is expected to be important for team performance, as it could affect how efficiently team members communicate – in terms of the 'push' and 'pull' of information – as well as affecting whether the team recognises the requirement to transition between explicit and implicit modes of coordination (e.g., Entin & Serfaty, 1999). It is also indicative of

the extent of monitoring within the team, which can in turn affect the likelihood that team members will recognize when backing up behaviours are required.

The Task Mutual Awareness Questionnaire asks team members to indicate the tasks which they and other team members were undertaking at particular points during mission execution. This is done retrospectively, by prompting team members to think back to particular events during a recently completed mission. For each event, team members report the tasks they were performing when the event occurred, and then report the tasks each of the other team members was performing during the same event. One version of the questionnaire includes a free-response format which requires that a subject-matter expert make judgments about similarities and differences between the responses given by different team members. Another version of the questionnaire asks team members to choose tasks from a pre-specified list. The Team Workload Awareness Questionnaire is basically an extension of the NASA-TLX workload scale (e.g., Hart & Staveland, 1988). In the TMA context, the TLX is applied as a three-part questionnaire. In the first part, participants report their own workload. In the second part, participants provide an estimate of the workload experienced by each of the other team members. In the third part, participants provide estimates of the workload of the team 'as a whole'. The Teamwork Awareness Questionnaire is a modified version of the teamwork behaviour scales described in Section 2.3 of this report. In the standard application of these scales, expert observers rate teamwork behaviours on seven-point BARS items after watching the team perform a mission. In the TMA application, the team members themselves are asked to retrospectively provide ratings of their own teamwork behaviours.

Some evidence exists to support the validity of the TMA measures, as well as their relationship to team outcomes (Entin & Entin, 2000; Macmillan et al., 2004). The ability of these tools to facilitate assessment of team members' understanding of states and processes *within* the team means the TMA approach may be a useful adjunct to assessment of Team SA when research or training goals are focused on the quality of the team's collective understanding of a dynamic situation.

4.4 Conceptual Issues with Team Situation Awareness

There is not yet consensus on how best to conceptualise team SA. Accordingly, there is not yet consensus on how best to operationalise and quantify team SA. Currently, the most commonly used methods to measure team SA are unsophisticated extensions of existing individual-level measurement techniques. The lack of theoretical and methodological consensus surrounding the construct means that careful consideration is required in (1) deciding if measurement of team SA is required, (2) determining the way in which team SA will be operationalised, and (3) generalising from the results of team SA assessment. The decision to assess team SA should be driven by the specific goals of each training or research project. The conceptualisation and operationalisation of team SA should be driven by the nature of the team, and of the task environment, that is under consideration.

5. Measuring Motivational Factors

Motivational factors⁸ influence both an individual's propensity to act, and their decision to act in a particular way under particular circumstances (Dick & Carey, 1996). As mentioned in the general introduction, the relationship between emergent states such as these and team performance is complex and reciprocal (Marks et al., 2001; Sosik, Avolio, & Kahai, 1997). Although they can be viewed as both inputs to and outputs of the teamwork cycle, they are probably best conceptualised as variables that mediate the effect of team processes (Ilgen et al., 2005). This section of the report provides a discussion of the antecedents, effects, and measurement of three emergent motivational states that have been consistently shown to affect team performance: collective efficacy, team cohesion, and trust.

5.1 Collective Efficacy

The notion of collective efficacy is based on Bandura's (1977) concept of self efficacy, which he defined as an individual's conviction that a behaviour or action can be executed. Consistent with the notion of the team mental model, collective efficacy differs from the self efficacy beliefs of individual team members in that it is a team-level property that emerges from the dynamic interaction among team members (Bandura, 2000; Kozlowski & Ilgen, 2006). Hence, collective efficacy refers to the team's shared belief in its ability to execute a course of action effectively (Guzzo, Yost, Campbell, & Shea, 1993), and it has been purported to influence how teams select courses of action, the effort teams will exert to achieve a result, and their persistence under threat of failure (Paskevich, Brawley, Dorsch, & Widmeyer, 1999).

The construct of collective efficacy bears much in common with the construct of group potency. Salas, Burke, Fowlkes, and Priest (2004) argued that the latter construct has a more task-generic emphasis. However, the two terms have been used interchangeably by some authors in the literature (e.g., Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995; Jex & Bliese, 1999; Zaccaro, Blair, Peterson, & Zazanis, 1995), and the task-specific versus task-generic conceptual distinction is not evident in the nature of items in inventories used to assess the two constructs (for an illustration see Salas et al., 2004). For the purposes of this paper, the constructs will be treated as synonymous.

5.1.1 Antecedents and Effects of Collective Efficacy

The major antecedents to collective efficacy are prior performance, leadership, availability of resources, and task structure. If a team has consistently performed well together in the past, they are more likely to believe they are capable of doing so in the future (Lindsley, Brass, & Thomas, 1995; Zaccaro et al., 1995). It is important to note that failure at a task does not necessarily reduce efficacy beliefs, but may lead to an adjustment in strategy which then increases collective efficacy for subsequent performance (Mesch, Farh and Podsakoff, 1989;

⁸ In the literature, various labels have been used to refer to these factors, including attitudes, values, and affective factors (e.g., Gagne, 1984; Kraiger, Ford, & Salas, 1993). We use the term *motivational factors* here as this best describes their role in affecting decisions to act and the propensity to act in particular ways.

Zaccaro et al., 1995). Collective efficacy can also be fostered by leaders who develop a cohesive environment, clarify role requirements and expectations, consult with the team on decisions, and who are achievement focussed (Zaccaro et al., 1995). Beliefs about a team's ability to succeed at a task are also influenced by perceptions of having the resources (e.g., time, expertise, personnel) necessary for the task (Shea & Guzzo, 1987), and by the manner in which the task is structured. When there is a high level of task interdependence and cohesion, there is more opportunity to observe and evaluate team member's skills; if evaluations are favourable, collective efficacy beliefs may increase (Shea & Guzzo, 1987; Sosik et al., 1997).

Collective efficacy has been shown to predict team effectiveness in both laboratory (Sosik et al., 1997) and field settings (Shea & Guzzo, 1987). Generally, high perceptions of efficacy are associated with improved task performance (Guzzo et al., 1993; Sosik et al., 1997). However, as is the case with team mental models, the relationship between collective efficacy and performance is not always straightforward, and can depend on factors specific to the particular task domain in question (Jung & Sosik, 1999, Sosik et al., 1997).

5.1.2 Measurement of Collective Efficacy

Paskevich et al. (1999) identified four distinct approaches to the measurement of collective efficacy. The first involves aggregating individuals' evaluations of their own self efficacy. This approach is most relevant when there is a high degree of heterogeneity in team roles, as it does not rely on knowledge of the performance of other team members. An example of an inventory relevant to this approach, adapted from DeShon, Kozlowski, Schmidt, Milner and Wiechmann (2004) by Best et al. (2007) for use in distributed mission training exercises, is provided in Table A1 of the Appendix A. Individuals respond on a 5-point scale ranging from strongly disagree (1) to strongly agree (5). DeShon et al. reported that factor analysis of responses on this scale yielded a one-factor solution with coefficient alphas ranging from .95 to .96 over three successive administrations.

A potential shortcoming of this approach is that it does not address the fact that the concepts of self efficacy and collective efficacy are fundamentally different: while individual team members may believe themselves to be capable, this does not equate to a belief in the capability of the team as a whole (Bandura, 2000). DeShon et al. (2004) assessed individual team members' beliefs in the team's efficacy by modifying the items in the individual-level inventory so that the team served as a referent, rather than the individual (shown in the bottom part of Table A1). This team-level version also appeared unidimensional, and demonstrated high internal consistency (coefficient alphas ranged from .96 to .97 over three administrations).

A third method of assessing collective efficacy reported by Paskevich et al. (1999) reflects a more holistic approach, and asks the team as a whole to provide a single rating of collective efficacy. In this method team members discuss their assessment, and form a consensual rating. It is considered most valid when there is a high level of interdependence and interaction among team members, which enables individuals access to the beliefs of the group. Like holistic approaches to measuring team knowledge (see Footnote 4 in Section 3.4), there are problems with such approaches to measuring collective efficacy, such as demand characteristics, pressure to conform to team norms, and the possibility of disproportionate influence by some team members (Bandura, 2000).

The fourth approach treats the individual as an informant and asks them to estimate the team's level of collective efficacy. This approach has been employed by Paskevich et al. (1999) in the assessment of collective efficacy in sporting teams, and contains items such as 'Our team's confidence that we have set realistic goals is _____. The measure employed by Paskevich et al. shows high internal consistency, with coefficient alphas ranging from .86 to .93. According to Paskevich et al. the second and fourth methods are the most theoretically and experimentally sound measures of collective efficacy (Paskevich et al., 1999). The choice of a particular approach depends primarily on the nature of the team, as the fourth measure assumes that team members could be reasonably expected to have access to the beliefs of the group (Paskevich et al., 1999). While this may be the case for co-located, highly interdependent teams, it may not be applicable to distributed or virtual teams (Cuevas, Fiore, Salas, & Bowers, 2004).

5.2 Team Cohesion

Team cohesion has been defined as "the resultant of all the forces which are acting on the members to stay in a group" (Festinger, 1950, p. 274). It is purported that there are three main components to team cohesion: interpersonal attraction, commitment to the task, and group pride (Beal, Cohen, Burke, & McLendon, 2003; Festinger, 1950; Mullen & Copper, 1994). Interpersonal attraction is a shared admiration or attachment to the members of the group. Task commitment refers to the extent to which there is a mutual commitment to the task, and the extent to which the task enables the group to achieve important goals. Group pride is concerned with the shared importance of being a team member, and is evident in the degree to which team members show affiliation with what the team represents and symbolises (Beal et al., 2003). Of the three, the strongest predictor of team performance is commitment to the task (Beal et al., 2003; Mullen & Copper, 1994).

According to Mullen and Cooper (1994) the consensus in the research is that there is a small but significant positive relationship between team cohesion and performance. The strength of this relationship depends on the moderating effects of team size, degree of interaction, level of group reality, and the type of cohesion that is under consideration – task versus interpersonal (Mullen & Cooper, 1994; Zaccaro, 1991). According to Mullen and Cooper (1994) team cohesion influences team performance through its effects on team coordination: high team cohesion improves coordination processes, which in turn improve performance on the task.

Team cohesion is closely related to collective orientation, which is characterised by a willingness to take into account input from fellow team members, and a belief that the goals of the team take precedence over those of the individual (Cannon-Bowers et al., 1995). Miles (2000) found that these two constructs were highly correlated, were related in a similar way to team outcomes, and were represented by a single underlying factor. Consequently, for the purposes of this paper collective orientation is considered to be part of the broader construct of team cohesion.

There are a number of different theoretical models of team cohesion. The two models of most relevance to the type of teams that are the focus of this report, and which have generated the most useful measures, are described below.

5.2.1 A Model of Military Team Cohesion

The degradation of unit identity and social alienation of troops during the end stages of the Vietnam War has been attributed in part to a lack of unit cohesion (Griffith, 1988). According to Griffith, this recognition acted as a catalyst for military research on team cohesion in combat units, and led to the implementation of new systems for managing the deployment of soldiers. In an attempt to increase the perception of belonging to a specific group, and therefore increase the unity and cohesion within that group, the units formed in training were maintained throughout deployment and assignment. In order to evaluate the effectiveness of this system, operational measures of military unit cohesion were developed (Griffith, 1988).

Griffith (1988) found that the quality of the relationships between teams of junior soldiers, and between junior soldiers and their leaders, was positively related to willingness to stay in the unit, morale, satisfaction with the Army, and perceptions of combat readiness. Subsequent research has identified specific types of cohesion existing within military teams. Siebold (2005) proposed three distinct types of cohesion: vertical, horizontal, and organisational. Vertical cohesion describes the relationships between subordinates and their leaders, horizontal cohesion describes the relationships between peers, and organisational cohesion refers to the relationship between unit members and the unit as a whole. This tripartite distinction has since received support within the domain of military psychology (Bliese & Halverson, 1996; Siebold, 2005; Siebold & Kelly, 1988).

Siebold and Kelly (1988) proposed that each of these three types of cohesion has two aspects: affective cohesion and instrumental cohesion. The affective aspect relates to emotions and reactions, whereas the instrumental aspect relates to the task. This division results in six distinct components of cohesion outlined in Figure 5.1. Horizontal Cohesion–Affective represents the feelings of trust and friendliness within a group, while Horizontal Cohesion–Instrumental refers to how well peers work together to achieve a task. Vertical Cohesion–Affective is concerned with the extent to which leaders consider the best interests of their subordinates, while Vertical Cohesion–Instrumental pertains to perceptions of the skills and ability of the leader to lead subordinates in training and combat. Organisational Cohesion–Affective is the extent to which individuals identify with the unit and its values, and feel pride in unit membership. Organisational Cohesion–Instrumental is the extent to which individuals attempt to achieve the goals of the unit, or organisation, in exchange for the unit facilitating the individual to attain their individual goals.

These six components of cohesion are assessed using the Combat Platoon Cohesion Questionnaire (CPCQ; Siebold & Kelly, 1988). This 73-item questionnaire demonstrates reasonable internal consistency and convergent validity properties, and takes about 50 minutes to administer. However, the CPCQ is highly specific to combat platoons, and its length limits its utility in cases where cohesion is not the question of primary theoretical or applied interest. Table A2 in the Appendix contains shorter, more broadly applicable affective and instrumental cohesions scales adapted from the CPCQ for use in research on training for Air Battle Management teams (Best et al., 2007).

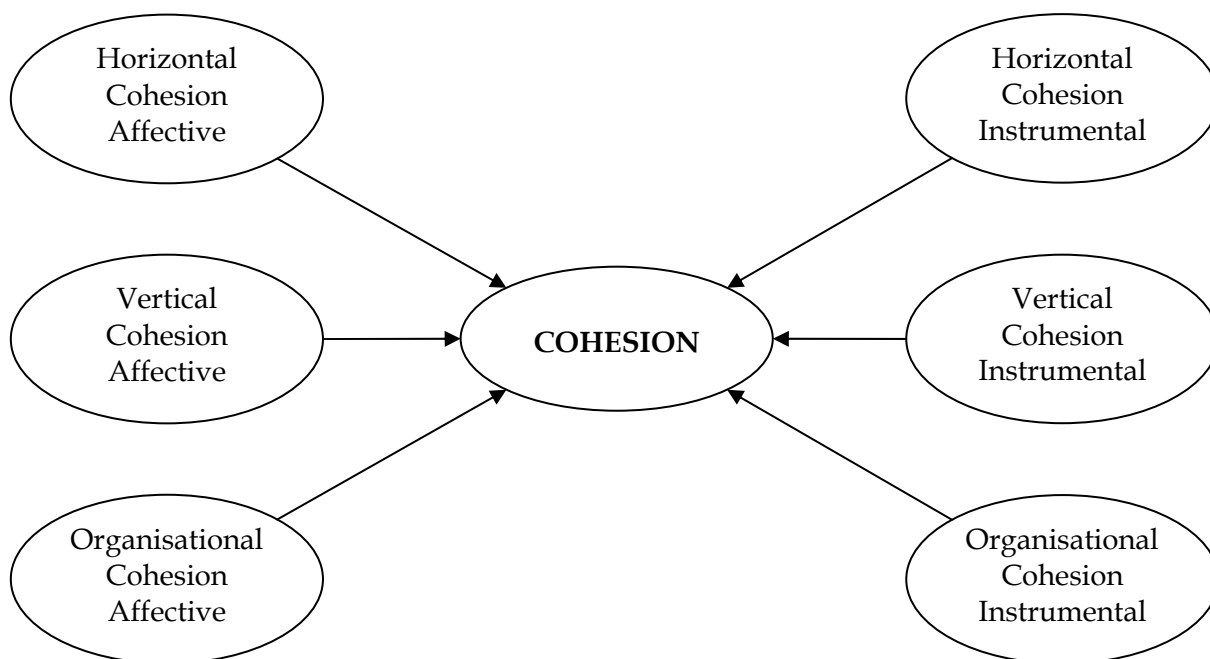


Figure 5.1 A taxonomy of team cohesion. Adapted from "Development of the combat platoon cohesion questionnaire," by G. L. Siebold and D. R. Kelly, 1988, ARI Technical Report 817, p. 3.

5.2.2 Carron's Hierarchical Model of Cohesion

Carron, Widmeyer, & Brawley (1985) argued that team cohesion can be separated into four distinct elements. Team cohesion (referred to as group cohesion in their model) is separated into a group component (group integration) and an individual component (individual attraction to group). Group integration is concerned with individuals' perceptions of the level of integration and familiarity within the team, and individual attraction to the group is concerned with members' personal commitment and association with the team. These components are further divided into task related and socially related components, creating the four components of cohesion depicted in Figure 5.2: Group integration-social (GI-S), Group Integration-Task (GI-T), Individual Attraction to the Group-Task (ATG-T), and Individual Attraction to the Group-Social (ATG-S). The distinction between task and social cohesion is supported by research which suggests that task cohesion and interpersonal cohesion show different relations with team performance (Zaccaro, 1991).

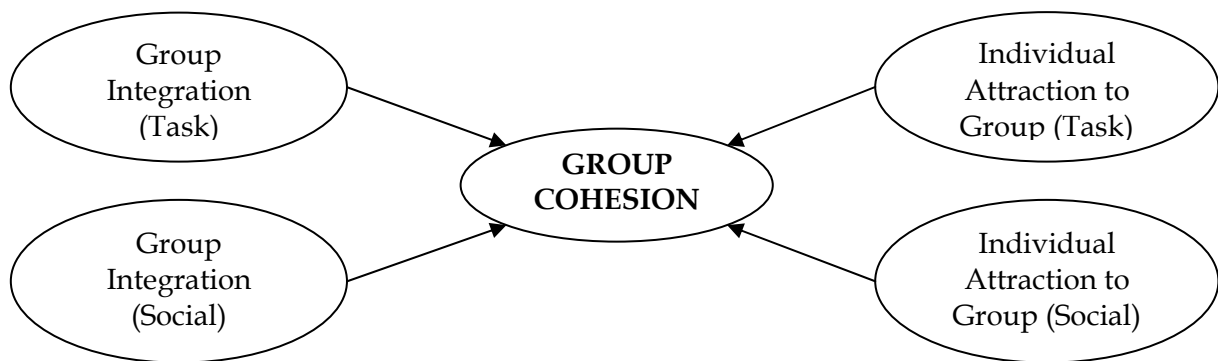


Figure 5.2. Hierarchical model of group cohesion. Adapted from "The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire," by A. V. Carron, W. N. Widmeyer, and L. R. Brawley, 1985, *Journal of Sport Psychology*, 7, p. 248.

Carron and colleagues (Carron et al., 1996; Carron et al., 1985) developed the self-report Group Environment Questionnaire (GEQ) to assess these four components of cohesion. The GEQ features 18 items rated on a 9-point Likert scale, and displays sound psychometric properties. The GEQ contains items such as 'our team would like to spend time together in the off season' (GI-S), 'our team is united in trying to reach its goals for performance' (GI-T), 'I'm not happy with the playing time I get' (ATG-T), and 'for me this team is one of the most important social groups to which I belong' (ATG-S). Although the GEQ was originally designed to be used in the context of sporting teams, it can be modified to suit other requirements (Dion, 2000).

5.3 Trust

Trust can be defined as "the shared perception by the majority of team members that individuals in the team will perform particular actions important to its members and that the individuals will recognise and protect the rights and interests of all the team members engaged in their joint endeavour" (Webber, 2002, p. 205). The concept of trust encompasses the notion of psychological safety (Edmondson, 1999), which is a shared belief that the team environment is safe for risk taking – for example, offering dissenting opinions or innovative ideas (Mayer, Davis, & Schoorman, 1995). Trust is particularly important in teams where there is a high degree of interdependence among team members, where there is a task that requires considerable collaboration, and where there is a need to take risks (Salas et al., 2004).

Trust encompasses both the propensity to trust and perceptions of trustworthiness (Costa, 2003; Mayer et al., 1995). Propensity to trust refers to a general willingness to trust others, and is a relatively stable personality characteristic that does not vary across trustees. Perceptions of trustworthiness involve evaluation of the motives and intentions of a team member, and can vary according to the trustee. Closely associated with the concept of trust are cooperative behaviour and a lack of monitoring behaviour; these variables account for a significant portion of the variance in the concept (Costa, 2003).

5.3.1 Antecedents and Effects of Trust

The simplest way to foster trust within a team is to form a team where the members already know and think favourably of one another. Forming a team that contains people who are known to be highly able, or who have previously worked together successfully, increases the likelihood that an individual will trust their team members from the outset (McAllister, 1995; Webber, 2002). These favourable conditions are not always available, and at some point team members are likely to face the need to take a risk without the reassurance of familiarity. This behaviour both necessitates and facilitates the development of trust: when a risk is taken a degree of trust is assumed by the risk taker; this then facilitates the development of trust in other team members (Costa, 2003). For this reason, risk taking - demonstrating a willingness to be vulnerable to team members - is seen to be a trust behaviour (Rousseau, Sitkin, Burt, & Camerer, 1998). If a favourable outcome ensues from this risk taking, the risk taker's sense of trust of is reinforced.

Webber (2002) suggested that greater commitment to the task will also foster increased trust within the team. If team members are committed to the task, and have a personal investment in the task's outcome, then perceptions of trust will increase. Commitment to the task can be achieved by perceptions of personal gain that result from successful completion of the task (Webber, 2002). For instance, if success in the task is related to career advancement, team members will feel more personally invested in the task. Webber argues that this sense of personal investment can also be achieved by developing overarching goals or missions for the task, which act to create a shared objective for team members.

Porter and Lilly (1996) found that while trust did not have a direct effect on performance, it had an indirect effect on performance through team process variables. These findings were supported by Dirks (1999), who found no evidence of a direct relationship between trust and task performance but found that trust moderated the relation between motivation, team processes, and performance. Additionally, Erdem, Ozen and Atsan (2003) found that the relation between trust and perceived task performance differed significantly between organisations. This study, however, did not take into account the potential confounding effect of team process variables, and it employed measures which, although possessing reasonable face validity, are yet to be the subject of convincing psychometric validation. The trust scale developed by Erdem et al. (2003), as employed in distributed mission training exercises by Best et al. (2007), is included in the Appendix as Table A3.

In apparent contradiction to these studies, Costa (2003) found that high levels of trust were directly and positively related to perceived task performance and team satisfaction. However, it should be noted that Costa measured trust as four distinct components, each of which related differently to performance and satisfaction. This differs from previous studies, which have operationalised trust as one factor and have investigated the effects of mediators and moderators of trust on performance. It would seem that when the effects of mediators are taken into account, the pattern of results found by Costa is consistent with the notion of emergent states as mediating (or moderating) variables in the teamwork cycle.

5.3.2 Measurement of Trust

Costa (2000) developed a measure of trust consistent with her notion of the concept, as described above. This measure assesses both the propensity to trust and perceptions of trustworthiness, as well as cooperative and monitoring behaviours. Examples of items relevant to these four dimensions are: 'people usually tell the truth, even when they know they would be better off lying' (propensity to trust); 'in my team some people have success by stepping on other people' (perceptions of trustworthiness); 'In my team we provide each other with timely information' (cooperative behaviours); and 'in my team people check whether others keep their promises' (monitoring behaviours). Responses to these items are made on a 7-point Likert scale (ranging from 'correct' to 'completely incorrect'). Costa's trust scales show adequate internal consistency, with alphas greater than .70.

McAllister (1995) developed an 11-item measure to assess affect- and cognition-based trust. According to McAllister, cognition-based trust depends on what the trustor believes are sound, evidence-based reasons for trusting – such as demonstrating behaviours that attest to the individual's trustworthiness. Affective-based trust, on the other hand, depends on the emotional ties and investment that form the relationship between the trustor and an individual (McAllister, 1995). The measure, displayed in Table A4 in the Appendix, contains six items that refer to cognition-based trust, and five that refer to affect-based trust. Respondents indicate the extent to which they agree or disagree with these statements on a 7-point likert scale. The cognition- and affect-based scales have been shown to possess high internal consistency (Cronbach's alphas of .91 and .89, respectively).

5.4 Applications and Conclusions

The research reviewed above suggests that emergent states such as collective efficacy, cohesion, and trust are associated with team performance. The clear implication is that by including these variables in measurement approaches for teamwork research or team training, there is the potential for improving team outcomes. The measures described above provide a means of gauging the existing level of these variables in a particular team, and evaluating the efficacy of interventions geared at improving them. It is also clear that the nature of the relationship between these emergent states and team performance is complex. Although they are most likely to act as moderators or mediators of team processes and inputs, further research is required to clarify the exact mechanisms by which they have an effect.

6. Conclusion

It may seem attractive to simply base inferences about team performance on observations of the outcomes of team tasks. However, this approach does not take into account the processes and emergent states which give rise to these outcomes, and is therefore of little utility to researchers striving to understand teams, or to training specialists striving to achieve improvements in teamwork. In contrast, the measurement of the behavioural, cognitive, and motivational constructs discussed in this report enable a more direct examination of the true determinants of team task outcomes. We hope that the material presented here will allow both researchers and training specialists to make informed judgments about how to tailor measurement approaches to their specific needs.

7. References

1. Alexander, A. L., & Wickens, C. D. (2005). Synthetic vision systems: Flightpath tracking, situation awareness, and visual scanning in an integrated hazard display. *Proceedings of the 13th International Symposium on Aviation Psychology*. Oklahoma City, OK: FAA.
2. Anderson, J. R. (1976). *Language, memory and thought*. Hillsdale, NJ: Erlbaum.
3. Annett, J. (1997). Analysing team skills. In R. Flin, E. Salas, M. Strub, & L. Martin (Eds.), *Decision Making Under Stress: Emerging Themes and Applications* (pp. 315-325). Aldershot: Ashgate.
4. Annett, J., Cunningham, D., & Mathias-Jones, P. (2000). A method for measuring team skills. *Ergonomics*, 43 (8), 1076-1094.
5. Austin, J. R. (2003). Transactive memory in organizational groups: The effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology* 88(5), 866-878.
6. Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, 84(2), 191-215.
7. Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science*, 9(3), 75-78.
8. Beal, D. J., Cohen, R. R., Burke, M. J., & McLendon, C. L. (2003). Cohesion and performance in groups: A meta-analytic clarification of construct relations. *Journal of Applied Psychology*, 88(6), 989-1004.
9. Best, C., Hasenbosch, S., Skinner, M., Crane, P., Burchat, E., Finch, M., Gehr, S. E., Kam, C., Shanahan, C., & Zamba, M. (2007). Exercise Pacific Link 2: Coalition distributed mission training. In *Proceedings of the SimTecT 2007Conference*. Brisbane, Australia: Simulation Industry Association of Australia.
10. Best, C., Sauvage, M.C., & Skinner, M. (in press). Training for Network-Enabled Air Operations: Competencies, Strategies and Technologies. To appear in R. Bolia & R. Taylor (Eds.), *Supporting Decision Effectiveness in Network-Centric Operations*. Dayton, OH: Wright-Patterson AFB.
11. Bolia, R. S., & Nelson, W. T. (2007). Characterizing team performance in network-centric operations: Philosophical and methodological issues. *Aviation, Space, and Environmental Medicine*, 78, B71-B76.
12. Bliese, P. D., & Halverson, R. R. (1996). Individual and nomothetic models of job stress: An examination of work hours, cohesion, and well-being *Journal of Applied Social Psychology*, 26(13), 1171-1189.

13. Brannick, M. T., Prince, A., Prince, C., & Salas, E. (1995). The measurement of team process. *Human Factors*, 37(3), 641-651
14. Brannick, M. T., Roach, R. M., & Salas, E. (1993). Understanding team performance: A multimethod study. *Human Performance*, 6(4), 287-308.
15. Breton, R., & Rousseau, R. (2003). Situation Awareness: A review of the concept and its measurement. In T. R. D. V. TR-2001-220 (Ed.). Valcartier, Canada: Defence Research and Development.
16. Brickman, B. J., Hettinger, L. J., Roe, M. M., Stautberg, D., Vidulich, M. A., Haas, M. W., & Shaw, R. S. (1995). An assessment of situation awareness in an air combat simulation: The global implicit measure approach.. In D. J. Garland & M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness*. Daytona Beach, FL: Embry Riddle Aeronautical University Press.
17. Burke, J. L., Murphy, R. R., Coovert, M. D., & Riddle, D. (2004). Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. *Human-Computer Interaction*, 19, 85-116.
18. Cannon-Bowers, J. A., & Salas, E. (2001). Reflections on shared cognition. *Journal of Organizational Behavior*, 22(2), 195-202.
19. Cannon-Bowers, J. E., Salas, E., and Converse, S. (1993). Shared mental models in expert team decision-making. In J. Castellan, (Ed.), *Individual and Group Decision-Making: Current Issues* (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum Associates.
20. 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 & Associates (Eds.), *Team Effectiveness and Decision Making in Organizations*. (pp. 333-380). San Francisco: Jossey-Bass.
21. Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The development of an instrument to assess cohesion in sport teams: The Group Environment Questionnaire. *Journal of Sport Psychology*, 7, 244-266.
22. Chi, M. (2006). Laboratory methods for assessing experts' and novices' knowledge. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. New York: Cambridge University Press.
23. Cohen, B., & Murphy, G. (1984). Models of concepts. *Cognitive Science*, 8, 27-58.
24. Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies*, 41, 801-849.
25. Cooke, N. J. (1999). Knowledge elicitation. In R. S. Nickerson & F. T. Durso (Eds.), *Handbook of applied cognition*. Chichester, NY: Wiley.

26. Cooke, N. J., Kiekel, P. A., Salas, E., Stout, R., Bowers, C., & Cannon-Bowers, J. (2003). Measuring team knowledge: A window to the cognitive underpinnings of team performance. *Group Dynamics: Theory, Research, and Practice*, 7(3), 179-199.
27. Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring team knowledge. *Human Factors*, 42(1), 151-173.
28. Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 83-106). Washington, DC: American Psychological Association.
29. Costa, A. C. (2003). Work team trust and effectiveness. *Personnel Review*, 32(5), 605-672.
30. Crane, P., Skinner, M., Best, C., Burchat, E., Gehr, E.S., Grabovac, M., Pongratic, H., Robbie, A., & Zamba, M. (2006). *Exercise Pacific Link: Coalition distributed mission training using low-cost communications*. Proceedings of the Australasian Simulation Technology and Training Conference (SimTecT), Melbourne, Australia.
31. Cuevas, H. M., Fiore, S. M., Salas, E., & Bowers, C. A. (2004). Virtual teams as sociotechnical systems. In S. H. Godar & S. P. Ferris (Eds.), *Virtual and Collaborative Teams* (pp. 1-19). Pennsylvania: Idea Group.
32. DeShon, R. P., Kozlowski, S. W. J., Schmidt, A. M., Milner, K. R., & Wiechmann, D. (2004). A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance. *Journal of Applied Psychology*, 89(6), 1035-1056.
33. Dick, W., & Carey, L. (1996). *The systematic design of instruction* (4th ed.). Glenview, IL: Scott, Foresman.
34. Dion, K. L. (2000). Group cohesion: From "Field of Forces" to multidimensional construct. *Group Dynamics: Theory, Research and Practice*, 4(1), 7-26.
35. Dirks, K. T. (1999). The effects of interpersonal trust on work group performance. *Journal of Applied Psychology*, 84(3), 445-455.
36. Dwyer, D., Fowlkes, J., Oser, R.L., Salas, E., & Lane, N.E. (1997). Team performance measurement in distributed environments: The TARGETS methodology. In M.T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods, and Applications* (pp. 137-153). Mahwah, NJ: Lawrence Erlbaum Associates.
37. Edmondson, A. (1999). Psychological safety and learning behaviour in work teams. *Administrative Science Quarterly*, 44(2), 350-383.
38. Edwards, B. D., Day, E. A., Arthur, W., Jr., & Bell, S. T. (2006). Relationships Among Team Ability Composition, Team Mental Models, and Team Performance. *Journal of Applied Psychology*, 91(3), 727-736.

39. 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.
40. Endsley, M. R. (1993). *Situation awareness and workload: Flip sides of the same coin*. Paper presented at the Proceedings of the Seventh International Symposium on Aviation Psychology, Columbus, OH.
41. Endsley, M. R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.
42. Endsley, M. R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*, 37(1), 64-85.
43. Endsley, M. R., & Jones, W. M. (1997). *Situation Awareness Information Dominance & Information Warfare* (Technical Report). Belmont, MA: Armstrong Laboratory, Wright-Patterson AFB.
44. Endsley, M. R. and Rodgers, M. D. (1996). Attention distribution and situation awareness in Air Traffic Control. *Proceedings of the 40th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica, CA.
45. Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, S. G. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. *Proceedings of the 42nd annual meeting of the Human Factors and Ergonomics Society*, Chicago, IL.
46. Entin, E., & Entin, E. (2001). Measures for evaluation of team processes and performance in experiments and exercises. *Proceedings of the 6th International Command and Control Research and Technology Symposium*, Annapolis, MD.
47. Entin, E.E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, 41(2), 312-325.
48. Erdem, F., Ozen, J., & Atsan, N. (2003). The relationship between trust and team performance. *Work Study*, 52(6/7), 337-340.
49. Ericsson, K. A. (2006). Protocol analysis and verbal thought: Concurrent verbalizations of thinking during experts' performance on representative tasks. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. New York: Cambridge University Press.
50. Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

51. Espinosa, J. A., Lerch, F. J., & Kraut, R. E. (2004). Explicit versus implicit coordination mechanisms and task dependencies: One size does not fit all. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 107-129). Washington, DC: American Psychological Association.
52. Festinger, L. (1950). Informal social communication. *Psychological Review*, 57, 271-282.
53. Fiore, S. M., & Schooler, J. W. (2004). Process mapping and shared cognition: Teamwork and the development of shared problem models. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 133-152). Washington, DC: American Psychological Association.
54. Fischer, U., McDonnell, L., & Orasanu, J. (2007). *Aviation, Space, and Environmental Medicine*, 78(5), B86-B95.
55. Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition* (2nd ed.). New York, NY: McGraw-Hill.
56. Flach, F. M. (1995). Situation Awareness: Proceed with caution. *Human Factors*, 37(1), 149-157.
57. Fowlkes, J., Dwyer, D., Oser, R., & Salas, E. (1998). Event-based approach to training. *The International Journal of Aviation Psychology*, 8(3), 209-221.
58. Fowlkes, J., Lane, N. E., Salas, E. S., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. *Military Psychology*, 6(1), 47-61.
59. French, H. T., Clarke, E., Pomeroy, D., Seymour, M., & Clark, C. R. (2003). *Psychophysiological Measures of Situation Awareness*. Paper presented at the Proceedings of Human Factors of Decision Making in Complex Systems, Dunblane, Scotland.
60. Gagne, R.M. (1984). Learning outcomes and their effects: Useful categories of human performance. *American Psychologist*, 39, 377-385.
61. Gonzalez, M. G., Burke, M. J., Santuzzi, A. M., & Bradley, J. C. (2003). The impact of group process variables on the effectiveness of distance collaboration groups. *Computers in Human Behavior*, 19, 629-648.
62. Gordon, S. E., Schmieder, K. A., & Gill, R. T. (1993). Conceptual graph analysis: Knowledge acquisition for instructional system design. *Human Factors*, 35(3).
63. Gorman, J. C., Cooke, N. J., Pederson, H. K., Connor, O. O., & DeJoode, J. A. (2005). Coordinated awareness of situation by teams (CAST): Measuring team situation awareness of a communication glitch. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.
64. Gorman, J. C., Cooke, N., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, 49, 1312-1325.

65. Gorman, J. C., Foltz, P. W., Kiekel, P. A., Martin, M. J., and Cooke, N. J. (2003). Evaluation of latent semantic analysis-based measures of team communications content. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society*, Denver, CO.
66. Graesser, A. C., Clark, L. F. (1985). *Structures and procedures of implicit knowledge*. Norwood, NJ: Ablex Publishing Corporation.
67. Graesser, A. C., & Goodman, S. M. (1985a). How to construct conceptual graph structures. In B. K. Britton & J. Black (Eds.), *Understanding expository text* (pp. 363-384). Hillsdale, NJ: Erlbaum.
68. Graesser, A. C., & Goodman, S. M. (1985b). Implicit knowledge, question answering, and the representation of expository text. In B. K. Britton & J. Black (Eds.), *Understanding expository text* (pp. 109-171). Hillsdale, NJ: Erlbaum.
69. Griffith, J. (1988). Measurement of group cohesion in U.S. Army units. *Basic and Applied Social Psychology*, 9(2), 149-171.
70. Gugerty, J. L. (1997). Situation awareness during driving: explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, 3(1), 42-66.
71. Guzzo, R., Yost, P., Campbell, R., & Shea, G. (1993). Potency in Groups: Articulating a construct. *British Journal of Social Psychology*, 32(1), 87-106.
72. Harbluk, J. L. and Noy, Y. I. (2002). *The impact of cognitive distraction of driver visual behaviour and vehicle control*. Road Safety Directorate and Motor Vehicle Regulation Directorate, Ergonomics Division, Canada, Moshe Eizenman, EL-MAR Inc.
73. Hart, S.G., & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. Hancock & N. Mishkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: Elsevier.
74. Helmreich, R. L., Foushee, H. C., Benson, R., & Russini, W. (1986). Cockpit resource management: Exploring the attitude - performance linkage. *Aviation, Space and Environmental Medicine*, 57, 1198-1200.
75. Hoffman, R. R., & Lintern, G. (2006). Eliciting and representing the knowledge of experts. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*. New York: Cambridge University Press.
76. Hoffman, R. R., Crandall, B., & Shadbolt, N. (1998). A case study in cognitive task analysis methodology: The Critical Decision Method for the elicitation of expert knowledge. *Applied Cognitive Psychology*, 6, 1-48.

77. Hogg, D. N., Folleso, K., Torralba, B. & Volden, F. S. (1995). Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. *Ergonomics*, 38, 2394-2413.
78. Hogg, D. N., Torralba, B. & Volden, F. S. (1993). *A situation awareness methodology for the evaluation of process control systems: studies of feasibility and the implication of use*, (1993-03-05). Storefjell, Norway: OECD Halden Reactor Project.
79. Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From Input-Process-Output models IMOI models. *Annual Review of Psychology*, 56, 517-543.
80. Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors*, 41, 1-4.
81. Jex, S. M., & Bliese, P. D. (1999). Efficacy Beliefs as a Moderator of the Impact of Work-Related Stressors: A Multilevel Study. *Journal of Applied Psychology*, 84(3), 349-361.
82. Jung, D. I., & Sosik, J. J. (1999). Effects of group characteristics on work group performance: A longitudinal investigation. *Group Dynamics: Theory, Research and Practice*, 3(4), 279-290.
83. Kaber, D. B., & Endsley, M. A. (1998). Team situation awareness for process control safety and performance. *Process Safety Progress*, 17, 43-48.
84. Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, and Cybernetics*, 19, 462-472.
85. Klimoski, R. J., & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management*, 20(2), 403-407.
86. Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, 7(3), 77-124.
87. Kraiger, K., Ford, K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, 78(2), 311-328.
88. Krauss, R. M., & Fussell, S. R. (1990). *Mutual knowledge and communicative effectiveness*. Paper presented at the Intellectual teamwork: Social and technological foundations of cooperative work, Hillsdale, NJ.
89. Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259-284.
90. Langan-Fox, J., Code, S., & Langfield-Smith, K. (2000). Team mental models: Techniques, methods, and analytic approaches. *Human Factors*, 42(2), 242-271.

91. Levine, J. M., Resnick, L. B., & Higgins, E. T. (1994). Social foundations of cognition. *Annual Review of Psychology*, 44, 585-612.
92. Lewis, K. (2003). Measuring transactive memory systems in the field: Scale development and validation. *Journal of Applied Psychology*, 88, 587-604.
93. Lim, B. C., & Klein, K. J. (2006). Team Mental Models and Team Performance: A Field Study of the Effects of Team Mental Model Similarity and Accuracy. *Journal of Organizational Behavior*, 27(4), 403-418.
94. Lindsley, D. H., Brass, D. J., & Thomas, J. B. (1995). Efficacy-performance spirals: A multilevel perspective. *Academy of Management Review*, 20(3), 645-678.
95. MacMillan, J., Entin, E., Hess, K., & Paley M. (2004). Measuring performance in a scaled world: Lessons learned from the distributed dynamic-decision making (DDD) synthetic team task. In S. Schiflett, L. Elliott, E. Salas, & M. Coovert (Eds.), *Scaled Worlds: Development, Validation and Applications* (pp. 154-180). Burlington, VT: Ashgate.
96. MacMillan, J., Paley, M.J., Entin, E.B., & Entin, E.E. (2004). Questionnaires for distributed assessment of team mutual awareness. In N.A. Stanton, A. Hedge, K. Brookhuis, E. Salas, & H. Hendrick (Eds.) *The Handbook of Human Factors and Ergonomic Methods* (pp. 51-1 – 51-9). Boca Raton, FL: CRC Press.
97. Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26(3), 356-376.
98. Marks, M. A., Sabella, M. J., Burke, C., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology*, 87(1), 3-13.
99. Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Cannon-Bowers, J. A., & Salas, E. (2005). Scaling the quality of teammates' mental models: Equifinality and normative comparisons. *Journal of Organizational Behavior*, 26(1), 37-56.
100. Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *Journal of Applied Psychology* 85(2), 273-283.
101. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, 20(3), 709-734.
102. McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organisations. *Academy of Management Journal*, 38(1), 24-59.
103. Mesch, D. J., Farh, J. L., & Podsakoff, P. M. (1989). *Effects of feedback sign on group goal setting, strategies, and performance: An emperical examination of some control theory hypotheses*. Paper presented at the Organizational Behavior Division at the annual metting of the Academy of Management, Washington, DC.

104. Miles, J. A. (2000). Relationships of collective orientation and cohesion to team outcomes. *Psychological Reports*, 86, 435-444.
105. Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
106. Militello, L. G., Kyne, M. M., Klein, G., Getchell, K., & Thorsden, M. (1999). A synthesised model of team performance. *International Journal of Cognitive Ergonomics*, 3(2), 131-158.
107. Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, 22(2), 89-106.
108. Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin*, 115(2), 210-227.
109. Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
110. Orasanu, J. M. (1994). Shared problem models and flight crew performance. In N. Johnston, N. McDonald & R. Fuller (Eds.), *Aviation psychology in practice* (pp. 255-285). Aldershot, England: Ashgate.
111. Orasanu, J. M. (1995). *Evaluating team situation awareness through communication*. Paper presented at the Proceedings of an international conference on experiment analysis and measurement of situation awareness, Daytona Beach, FL.
112. Paskevich, D. M., Brawley, L. R., Dorsch, K. D., & Widmeyer, W. N. (1999). Relationship between collective efficacy and team cohesion: Conceptual and measurement issues. *Group Dynamics: Theory, Research and Practice*, 3(3), 210-222.
113. Pew, R. (1995). *The state of situation awareness measurement: Circa 1995*. Paper presented at the Proceedings of an international conference on experimental analysis and measurement of situation awareness, Daytona Beach, FL.
114. Pew, R. W. (2000). The state of situation awareness measurement: Heading toward the next century. In M. R. Endsley & D. J. Garlands (Eds.), *Situation Awareness Analysis and Measurement* (pp. 33-51). Manwah, NJ: Lawrence Erlbaum Associates.
115. Porter, T. W., & Lilly, B. S. (1996). The effects of conflict, trust and task commitment on project team performance. *International Journal of Conflict Management*, 7(4), 361-376.
116. Prince, C., Ellis, E., Brannick, M.T., & Salas, E. (2007). Measurement of team situation awareness in low experience level aviators. *The International Journal of Aviation Psychology*, 17(1), 41-57.

117. Proctor, M. D., Panko, M., & Donovan, S. J. (2004). Considerations for training team situation awareness and task performance through PC-gamer simulated multiship helicopter operations. *International Journal of Aviation Psychology*, 14(2), 191-205.
118. Regal, D. M., Rogers, W. H. & Boucek, G. P. Jr. (1988). *Situational awareness in the commercial flight deck: definition, measurement, and enhancement*. SAE Technical Paper Series: Aerospace Technological Conference and Exposition, Anaheim, CA.
119. Rentsch, J. R., & Woehr, D. J. (2004). Quantifying congruence in cognition: Social relations modeling and team member schema similarity. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 11-31). Washington, DC: American Psychological Association.
120. Robertson, M., & Endsley, M. R. (1995). The role of crew resource managements (CRM) in achieving team situational awareness. *Proceedings of the 21st Conference of the European Association for Aviation Psychology*, London, UK.
121. Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 8, 382-439.
122. Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100(3), 349-363.
123. Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *The Academy of Management Review*, 23(3), 393-404.
124. Rousseau, V., Aubé, C., Savoie, A. (2006). Teamwork behaviours: A review and an integration of frameworks. *Small Group Research*, 37(5), 540-570.
125. Salas, E., Burke, C. S., Fowlkes, J. E., & Priest, H. A. (2004). On measuring teamwork skills. In J. C. Thomas & M. Hersen (Eds.), *Comprehensive Handbook of Psychological Assessment*. (Vol. Vol. 4: Industrial and Organizational Assessment, pp. 427-442). Hoboken, NJ: Wiley.
126. Salas, E., & Cannon-Bowers, J. A. (1997). Methods, tools, and strategies for team training. In M. A. Quinones & A. Ehrenstein (Eds). *Training for a rapidly changing workplace: Applications of psychological research*. (pp. 249-279). Washington, DC, US: American Psychological Association.
127. Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. J. (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.
128. 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.
129. Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a “big five” in teamwork? *Small Group Research*, 36(5), 555-599.

130. Sarter, N. B., & Woods, D. D. (1991). Situation Awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
131. Sarter, N. B., & Woods, D. D. (1997). Teamplay with a powerful and independent agent: A corpus of operational experiences and automation surprises on the airbus A320. *Human Factors*, 39, 553-569.
132. Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Hillsdale, NJ: Erlbaum.
133. Schneider, D. J. (1991). Social cognition. *Annual Review of Psychology*, 42, 527-561.
134. Schraagen, J. M. (2006). Task analysis. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 185-201). London, UK: Cambridge University Press.
135. Sebok, A. (2000). Team performance in process control: influences of interface design and staffing levels. *Ergonomics* 43(8), 1210-1236.
136. Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations* (pp. 5/1-5/8). France: NATO-AGARD: Neuilly Sur Seine.
137. Serfaty, D., Entin, E., & Johnston, J.H. (1998). Team coordination training. In J.A. Cannon-Bowers & E. Salas (Eds.), *Making Decisions Under Stress: Implications for individual and team training* (pp. 221-245). Washington, DC: APA Press.
138. Shea, G. P., & Guzzo, R. A. (1987). Group effectiveness: What really matters? *Sloan Management Review*, 28(3), 25-31.
139. Siebold, G. L. (2005). Military group cohesion. In T. Britt, A. Adler & C. A. Castro (Eds.), *Military life: The psychology of serving in peace and combat*. (Vol. 1, pp. 185-201). Westport, Conn.: Praeger Security International.
140. Siebold, G. L., & Kelly, D. R. (1988). *Development of the combat platoon cohesion questionnaire* (No. ARI Technical Report 817): U.S. Army, Research Institute for the Behavioral and Social Sciences.
141. Smith, E. R. (1998). Mental representation and memory. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology* (4th ed.). New York: Oxford University Press.
142. Smith-Jentsch, K.A., Johnston, J.H., & Payne, S.C. (1998). Measuring team-related expertise in complex environments. In J. Cannon-Bowers & E. Salas (Eds.), *Making Decisions Under Stress: Implications for individual and team training* (pp. 61-87). Washington, DC: APA Press.

143. Smith-Jentsch, K.A., Zeisig, R.L., Acton, B., & McPherson, J.A. (1998). Team dimensional training: A strategy for guided self-correction. In J. Cannon-Bowers & E. Salas (Eds.), *Making Decisions Under Stress: Implications for individual and team training* (pp. 271-298). Washington, DC: APA Press.
144. Sosik, J. J., Avolio, B. J., & Kahai, S. S. (1997). Effects of leadership style and anonymity on group potency and effectiveness in a group decision support system environment. *Journal of Applied Psychology*, 82(1), 89-103.
145. Stanton, N. A., Chambers, P. R. G., & Piggott, J. (2001). Situational Awareness and Safety. *Safety Science*, 39, 189-204.
146. Stout, R., Salas, E., & Fowlkes, J.E. (1997). Enhancing teamwork in complex environments through team training. *Group Dynamics: Theory, Research, and Practice*, 1(2), 169-182.
147. Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors*, 41(1), 61-71.
148. Suppes, P., Pavel, M., & Falmagne, J. (1994). Representations and models in psychology. *Annual Review of Psychology*, 45 517-544.
149. Truitt, T. R., & Ahlstrom, V. (2000). *Situation awareness in airway facilities: Replacement of maintenance control centers with operations control centers*. (DOT/FAA/CT-TN00/09). Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.
150. Vidulich, M. A. (1995). The role of scope as a feature of situation awareness. In D. J. Garland & M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness*. Daytona Beach, FL: Embry Riddle Aeronautical University Press.
151. Venturino, M., Hamilton, W. L., & Dvorchak, S. T. (1990). Performance-based measures of merit for tactical situation awareness. *Situational Awareness in Aerospace Operations*, 5, 23-53.
152. Volpe, C.E., Cannon-Bowers, J.A., Salas, E., & Spector, P.E. (1996). The impact of cross-training on team functioning: An empirical investigation. *Human Factors*, 38(1), 87-100.
153. Webber, S. S. (2002). Leadership and trust facilitating cross-functional team success. *The Journal of Management Development*, 21(3/4), 201-214.
154. Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behaviour* (pp. 185-205). New York: Springer-Verlag.
155. Wegner, D. M., Erber, R., & Raymond, P. (1991). Transactive memory in close relationships. *Journal of Personality and Social Psychology*, 61(6), 923-929.

156. Wellens, A. R. (1993). Group situation awareness and distributed decision making from military to civil applications. In N. J. C. Jr (Ed.), *Individual and group decision making: Current issues* (pp. 267-291). Mahwah, NJ: Lawrence Erlbaum Associates.
157. Wickens, C. D., McCarley, J. S., Alexander, A. L., Thomas, L. C., Ambinder, M., & Zheng, S. (2005). *Attention-Situation Awareness (A-SA) Model of Pilot Error* (Technical Report AHFD-04-15/NASA-04-5). Moffett Field, CA: NASA Ames Research Center.
158. Zaccaro, S. J. (1991). Nonequivalent associations between forms of cohesiveness and group-related outcomes: Evidence for multidimensionality. *Journal of Social Psychology*, 131, 387-399.
159. Zaccaro, S. J., Blair, V., Peterson, C., & Zazanis, M. (1995). Collective efficacy. In J. E. Maddux (Ed.), *Self-Efficacy, Adaptation, and Adjustment: Theory, Research, and Application*. (pp. 305-328). New York: Plenum Press.

Appendix A: Scales for the Measurement of Motivational Factors

Table A1. Scales for assessing collective efficacy by aggregating individual-referenced and team-referenced responses

Individual-level scale	
1.	I can meet the challenges of the task.
2.	I understand how information cues are related to decisions.
3.	I can deal with decisions under ambiguous conditions.
4.	I can manage the requirements of the task.
5.	I will cope if the workload is increased.
6.	I will cope if the task becomes more complex.
7.	I can develop methods to handle changing aspects of the task.
8.	I can cope with different components of the task competing for my time.
Team-level scale	
1.	This team can meet the challenges of the task.
2.	This team understands how information cues are related to decisions.
3.	This team can deal with decisions under ambiguous conditions.
4.	This team can manage the requirements of the task.
5.	This team will cope if the workload is increased.
6.	This team will cope if the task becomes more complex.
7.	This team can develop methods to handle changing aspects of the task.
8.	This team can cope with different components of the task competing for its time.

Note: Individuals respond on a 5-point scale ranging from strongly disagree (1) to strongly agree (5). Adapted from "A multiple-goal, multilevel model of feedback effects on the regulation of individual and team performance," by R. P. DeShon, S. W. Kozlowski, A. M. Schmidt, K. R. Milner, and D. Wiechmann, 2004, *Journal of Applied Psychology*, 89, p. 1043.

Table A2: Scales for measuring affective and instrumental cohesion

Affective cohesion	
1.	In this team the members really care about what happens to each other.
2.	Members of this team can trust one another.
3.	Members of this team feel very close to each other.
4.	Members of this team like being in this team.
5.	Members of this team really respect one another.
6.	Members of this team like one another.
Instrumental cohesion	
1.	The members of my team make each other feel like doing a good job.
2.	The members of my team work well together.
3.	The members of my team help each other to get the job done.
4.	The members of my team encourage each other to succeed when performing the task.
5.	The members of my team work hard to get things done.
6.	The members of my team pull together and share the load while performing the task.

Note: Individuals respond on a 7-point scale ranging from disagree strongly (1) to agree strongly (7). Adapted from "Development of the combat platoon cohesion questionnaire," by G. L. Siebold and D. R. Kelly, 1988, ARI Technical Report 817, p. 3.

Table A3: Erdem, Ozen and Atsan's (2003) measure of trust

1.	Team members fulfil their undertakings successfully.
2.	Team members have the necessary qualifications for effective team performance.
3.	I can trust the expertise of team members.
4.	I can get help easily from team members.
5.	Team members share all sources with other members at all times.
6.	Team members encourage each other to introduce difference ideas and suggestions.
7.	Team members respect each other's emotions and ideas.

Note: The authors did not provide information about the rating scale they employed. Adapted from "The relationship between trust and team performance," by F. Erdem, J. Ozen, and N. Atsan, 2003, *Work Study*, 52, p. 339.

Table A4: McAllister's (1995) measure of affect- and cognition- based trust

Affect-based trust	
1.	We have a sharing relationship. We can both freely share our ideas, feelings and hopes.
2.	I can talk freely to this individual about difficulties I am having at work and know that (s)he will want to listen.
3.	We would both feel a sense of loss if one of us was transferred and we could no longer work together.
4.	If I shared my problems with this person, I know (s)he would respond constructively and caringly.
5.	I would have to say that we have both made considerable emotional investments in our working relationship.
Cognition-based trust	
1.	This person approaches his/her job with professionalism and dedication.
2.	Given this person's track record, I see no reason to doubt his/her competence and preparation for the job.
3.	I can rely on this person not to make my job more difficult by careless work.
4.	Most people, even those who aren't close friends of this individual, trust and respect him/her as a coworker.
5.	Other work associates of mine who must interact with this individual consider him/her to be trustworthy.
6. *	If people knew more about this individual and his/her background, they would be more concerned and monitor his/her performance more closely.

Note: Individuals respond on a 7-point scale ranging from strongly agree (1) to strongly disagree (7). Adapted from "Affect- and cognition-based trust as foundations for interpersonal cooperation in organisations," by D. J. McAllister, 1995, *Academy of Management Journal*, 38, p. 37.

* Denotes item should be reverse scored.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA					
				1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)	
2. TITLE Measurement of the Behavioural, Cognitive, and Motivational Factors Underlying Team Performance			3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)		
4. AUTHOR(S) Christopher Shanahan, Christopher Best, Melanie Finch and Christopher Sutton			5. CORPORATE AUTHOR Defence Science and Technology Organisation 506 Lorimer Street Fishermans Bend Victoria 3207 Australia		
6a. DSTO NUMBER DSTO-RR-0328		6b. AR NUMBER AR-013-933		7. DOCUMENT DATE June 2007	
8. FILE NUMBER 2007/1026887/1		9. TASK NUMBER 04/236		10. TASK SPONSOR DGAD	
11. NO. OF PAGES 71		12. NO. OF REFERENCES 159			
13. URL on the World Wide Web http://www.dsto.defence.gov.au/corporate/reports/DSTO-RR-0328.pdf				14. RELEASE AUTHORITY Chief, Air Operations Division	
15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT <i>Approved for public release</i> OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111					
16. DELIBERATE ANNOUNCEMENT No Limitations					
17. CITATION IN OTHER DOCUMENTS Yes					
18. DSTO RESEARCH LIBRARY THESAURUS http://web-vic.dsto.defence.gov.au/workareas/library/resources/dsto_thesaurus.htm teamwork, measurement, behaviour, and cognition					
19. ABSTRACT In the military, the use of teams to perform complex work functions is pervasive. Team performance is therefore critical, across a range of work contexts, for fulfilling operational requirements. To understand team performance, it is necessary to measure team performance. Valid and reliable measurement of the variables underlying effective team performance is fundamental to both team training and team training research, as it enables the determination of the strengths and weaknesses of teams, the design of targeted training programs, the provision of constructive feedback, and the evaluation of training effectiveness. This report presents, for the first time in a single resource, an overview of both the variables underlying effective team performance and the ways in which these variables have been operationalised and measured in the research literature.					