Ergonomics
Vol. 49, Nos. 12–13, 10–22 October 2006, 1312–1325

Measuring team situation awareness in decentralized command and control environments

JAMIE C. GORMAN*†§, NANCY J. COOKE‡§ and JENNIFER L. WINNER‡§

†Department of Psychology, New Mexico State University, Las Cruces, NM, USA
‡Department of Applied Psychology, Arizona State University, Mesa, AZ, USA
§Cognitive Engineering Research Institute, Mesa, AZ, USA

Decentralized command and control settings like those found in the military are rife with complexity and change. These settings typically involve dozens, if not hundreds to thousands, of heterogeneous players coordinating in a distributed fashion in a dynamically networked battlefield laden with sensor data, intelligence reports, communications, and plans emanating from many different perspectives. Consider the concept of team situation awareness in this setting. What does it mean for a team to be aware of a situation or, more importantly, of a critical change in a situation? Is it sufficient or necessary for all individuals on the team to be independently aware? Or is there some more holistic awareness that emerges as team members interact? We re-examine the concept of team situation awareness in decentralized systems beyond an individual-oriented knowledge-based construct by considering it as a team interaction-based phenomenon. A theoretical framework for a process-based measure called 'coordinated awareness of situations by teams' is outlined.

Keywords: Teams; Situation awareness; Command and control; Team cognition

1. General introduction

Since the 1980s the US government has placed an emphasis on speed, flexibility, and manoeuvre warfare in its force structure. In this regard, US service doctrine endorses highly malleable and adaptive decentralized command and control systems, concentrating less on centralized command and control architectures (Franz 2004). In decentralized systems there is no single central executive or leader directing every aspect of the battlefield, but rather responsibilities are distributed, culminating in an emergent coordination structure based on input from many different perspectives of the global terrain. This functionality comprises a general organizational strategy not limited to

*Corresponding author. Email: jagorman@nmsu.edu

Ergonomics
ISSN 0014-0139 print/ISSN 1366-5847 online
http://www.tandf.co.uk/journals
DOI: 10.1080/00140130600612788

military command and control systems, but applicable over a wide range of complex tasks, including emergency operations, business and academic domains, and distributed medical tasks (e.g. telemedicine).

The utility of decentralized team command and control resides in the concept of dispersing resources (individuals and their competencies) across the tactical environment of operations in order to carry out a strategic high-level function in a flexible and adaptive manner. A decentralized command and control (DC2) environment is a complex environment which requires coordination among a team of heterogeneously skilled operators in order to be effectively perceived and acted upon. Team DC2 environments can be quite large, involving hundreds of operators (e.g. network-centric warfare), or quite small, involving an ad hoc team of three or four (e.g. a distributed surgical team). In terms of team cognition, DC2 environments can be envisaged as a network of task elements distributed among operators according to a heterogeneous division of labour (Cooke and Gorman, in press). Task elements are coupled in the network such that no one operator must be cognizant of the DC2 environment as a whole, but rather operators are responsible for a simpler local environment, delineated by a subset of task elements.

By distributing responsibility, DC2 systems promise 'shared situation awareness' of the tactical environment (Hansen 2004); however, little is known about how to measure the fulfilment of this design principle in DC2 environments. In this paper we will argue that there are unexpected properties of DC2 environments which are not wholly contained in the individual knowledge properties of DC2 operators, i.e. functional properties that cannot be measured as a sum of each team member's individual awareness. Specifically, we will introduce a theoretical framework in which functional properties of DC2 teams which emerge from the dynamic interplay among networked task elements, through the interactions of team members, underlie the phenomenon of team situation awareness (TSA) in DC2 environments. In this paper we make use of this framework in introducing a theoretical and practical basis for measuring TSA called coordinated awareness of situations by teams (CAST).

1.1. *Team-level properties and phenomena*

Deterministic laws relating the micro-level of individual team member (operator) properties to the macro-level of team properties have yet to be discovered. In team research, this indeterminism has been attributed to deviation from a normative model of team process (Steiner 1972, Hinsz 1999), or alternatively a component model has been proposed with team-member properties causing team-level properties through a mediating team process factor (e.g. group interaction processes, Hackman 1987). Thus the component model suggests that team-level properties are not directly determined by team-member properties, but that team interaction processes are a ripe source of variance specific to teams. Because of this unique source of interaction-based variance, approaches that aggregate across the properties of individual team members, in order to determine a team-level property (e.g. Rentsch and Klimoski 2001), may be insufficient for measuring team-level phenomena (Cooke and Gorman 2006). Specifically, a team member cannot be the efficient cause of a team-level phenomenon because team-level properties are in large part the result of team-member interactions. For example, taken individually a team member is not heterogeneous, but still a team can exhibit this property when team members interact.

Extending Hackman's (1987) component model, we would further suggest that team members and the team coexist in a circular relation which simultaneously defines individual operators as 'team members' while driving the concept of 'team'. This circular model suggests that because team members are to some degree shaped by their participation on the team, the synthesis of team-level properties cannot be analytically reduced to the 'quanta' of operator properties, interaction processes, and team-level properties (see also Ilgen *et al.* 2005). Rather, we would argue that each of these levels of analysis is tightly coupled within the functional dynamics of team-level phenomena. We assume that a tight coupling between process and properties is especially salient in DC2 systems, as these systems are predicated on integrating multiple perspectives, and constitutes something more than an aggregate of the properties of DC2 team members. Specifically, we propose that team-level phenomena include team coordination processes and that these processes help shape the perceptual and action capabilities of highly interdependent team members. In this regard, we have centred our TSA measurement efforts on concepts such as 'coordinated perceptions' and 'coordinated actions'.

1.2. *Team situation awareness*

Historically, the aviation community is responsible for introducing the research community to a pervasive phenomenon of tactical flight operations called situation awareness (SA). However, in the human factors/ergonomics research community, there has been confusion about how to define this phenomenon (Durso and Gronlund 1999). Many SA researchers have agreed in principle on a three-part definition of SA: 'the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future' (Endsley 1988, p. 97). In this regard, SA can be considered either a cognitive product or a process (Smith and Hancock 1995); however, in practice it is most often treated as its own product (discussed by Flach 1995). On the contrary, we believe that SA should be viewed as a continuous perception–action process, in which ongoing activity plays an integral role in what there is to be perceived (Gibson 1998) during adaptation to an external constraint (Smith and Hancock 1995, Sandom 2001). That is, in highly dynamic situations there is often no time to reflect on the situation; action must guide perception and vice versa. Regarding combat pilots, for instance, we presume that perception and action during evasive manoeuvres in response to external threats, as a psychological phenomenon, involve pre-reflective cognitive processes. Accordingly, we believe that SA is best characterized as a pre-reflective process of adaptation, rather than a knowledge-based 'comprehension of meaning' (Endsley 1988, p. 97).

There is also a wide divergence of opinion concerning what measurement techniques best capture the essence of SA (Durso and Gronlund 1999). In accordance with the knowledge-based view, some SA researchers have sought to tie SA to knowledge structure by probing the memory retrieval processes of operators (e.g. long-term working memory (Sohn and Doane 2004)). Related to this method, many researchers have sought to measure SA using query methods, in which operators are probed during a task simulation for their dynamic understanding of various elements of the task environment (Endsley 1995). Also in line with the knowledge-based view, self-report or subjective measures (Taylor 1990) probe operators' retrospective assessments of their own SA. In contrast with knowledge-based approaches, some SA researchers have attempted to measure SA using implicit performance-based measures. These measures essentially assume that some tasks (or subtasks) require good SA to be performed well; if the operator has poor SA the task will be performed poorly, and thus a measure of SA is thought to be implicit in task performance (Wickens 1996). CAST is most similar to performance-based measurement, in that we expect teams with good SA to perform well. However, performance is always taking place, whether the situation is unanticipated or routine, and CAST focuses primarily on the coordinated processes involved in a team adapting to non-routine situational constraints.

Although adding the word 'team' to SA complicates matters, many researchers have approached measuring this seemingly more complicated phenomenon by aggregating the SA of the individual team members. This approach typically involves probing the projective mental states of team members during task performance (e.g. by querying them) in order to infer the knowledge, or model, of the situation and then aggregating to the team level by averaging, summing, or assessing the degree of overlap across individual mental states (e.g. a shared mental model) (Artman 1999, 2000, Fowlkes *et al.* 2000, Dekker 2000, Rasker *et al.* 2000, Cooke *et al.* 2001). The basis of this type of measure is that, as a cognitive product, SA 'exists only in the cognition of the human mind' (Bolstad and Endsley 2003, p. 369), and TSA is defined as 'the degree to which every team member possesses the SA [product] required for his/her job' (Endsley and Jones 1997). These SA requirements include 'shared SA requirements', such that each DC2 team member is aware of his/her independent task requirements as well as aspects of DC2 system coordination, i.e. their overlapping requirements (Endsley and Jones 1997). Summarizing this approach, it proposes that we measure TSA by probing team members' dynamic knowledge of the DC2 environment by querying them individually about their respective SA requirements and then summing the accuracy of their responses. CAST proposes a different approach to TSA measurement.

We assume that team DC2 environments, like other team environments, are inherently dynamic (Salas *et al.* 1992). If we further assume that team coordination changes over time in a dynamic task environment, then team coordination is always changing, ideally consistent with the dynamics of the task environment, in this case tracing out a 'trajectory' of DC2 operations. Drawing on the classic definition of team tasks provided by Salas *et al.* (1992), this trajectory is a continuous path, ideally focused on the common and valued goal. For example, 'health care' would be a common and valued goal for a medical DC2 team.

From our perspective however, measurement of TSA involves an additional layer of change in DC2 environments, in addition to the common and valued trajectory described by Salas *et al.* (1992). Not only are DC2 environments dynamic, but often these dynamics themselves can undergo change. The phenomenon of TSA corresponds to how an operational trajectory which spontaneously deviates from 'common and valued' is handled by a team through coordination processes in an attempt to maintain the common and valued integrity of the DC2 system as a whole; cf. 'it is imperative that pilots process and react quickly and appropriately to *unanticipated* events occurring beyond the cockpit' (Proctor *et al.* 2004, p. 192, emphasis added). In team DC2 this entails that a spontaneous 'change in changing coordination', in the light of events in the DC2 environment which could cause a DC2 team to deviate from its common and valued trajectory, should be assessed in terms of TSA. Therefore we believe that, when measuring TSA, it is not enough to query, observe, or record performance deficiencies from a DC2 team during conventional task performance. Rather, the concept of changes in the DC2 environment which push the DC2 system away from a common and valued trajectory would really put TSA to the test. For instance, in our unmanned aerial vehicle simulation (Cooke and Shope 2004, 2005) we characterized the process of team cognition using the metaphor of an inverted pendulum, an inherently unstable physical system which requires continuous feedback and control in order to balance it straight up.

Similarly, adaptive team cognition requires continuous coordination among team members in order to achieve a common and valued goal. In terms of TSA measurement we would like to flick the top of the team's 'pendulum' in order to measure their spontaneous 'balancing' response.

CAST measurement can address this using 'roadblocks'. In sum, a DC2 system is apt to follow a trajectory in accordance with the common and valued goals of the team, but if we put a roadblock in that path we can gain insight into the team's coordinated processes of perception and action, or TSA, as a path is navigated around the roadblock through teamwork. This is the basic idea behind CAST measurement.

1.3. *Practical differences between CAST and knowledge-based measures*

A 'mental model' (Craik 1943) is a dynamic knowledge-based representation in the mind that can be used to make predictions about the world. By creating shared expectations, some researchers have suggested that the development of a shared mental model (SMM) facilitates the development of good TSA (Bolstad and Endsley 1999, Prince and Salas 2000, Cooke *et al.* 2001). In terms of TSA, a SMM is analogous to the space 'where . . . SA requirements overlap' (Endsley and Jones 1997, p. 36ff). Overlapping SA requirements generally involve some high-level aspect of the DC2 environment, such as 'teamwork' knowledge regarding who talks to whom and when. From a knowledge-based perspective, in practice a TSA researcher would need to identify individual SA requirements 'in the cognition of the human mind' (Bolstad and Endsley 2003, p. 369) which overlap in the DC2 environment, in addition to heterogeneous SA requirements that do not overlap (Endsley *et al.* 1999). This can be accomplished by conducting a team cognitive task analysis (Seamster *et al.* 1997) of the DC2 environment in order to identify the operators' heterogeneous and overlapping SA requirements. Ideally, mental models and SMMs of the DC2 environment encompass these two types of SA requirements.

The most developed method for measuring knowledge-based SA in the human factors/ ergonomics literature is the situation awareness global assessment technique (SAGAT) (Endsley 1995, 2000). In SAGAT, the TSA researcher 'freezes' the task environment during team task performance in order to probe individual operators with queries concerning current or future states of their heterogeneous and overlapping SA requirements (Bolstad and Endsley 2003). Measuring TSA then proceeds by assessing the degree to which each operator accurately responds to SA queries involving overlapping SA requirements in the DC2 environment, in addition to queries concerning heterogeneous SA requirements. The more team members with accurate query responses, the higher is the TSA (Bolstad and Endsley 2003). Query accuracy is one of several allied criteria which can be used (Cooke *et al.* 2001).

While we do not posit that a query-based approach is inappropriate for assessing knowledge-based constructs, we believe that there are validity issues involved in employing such an approach for measuring highly dynamic phenomena such as SA. For example, if the goal is to evaluate team members regarding their independent levels of learning, knowledge elicitation techniques of various types are highly appropriate (Cooke 1994, 1999, Connor *et al.* 2004). However if the goal is to measure the pre-reflective cognitive processes underlying the adaptive ongoing awareness of an operator, knowledge elicitation is an indirect form of measurement. That is, to the extent that SA (and TSA) involves some degree of ongoing pre-reflective awareness, a proposition with which we assume many SA researchers would agree, knowledge elicitation may be, at best, an indirect form of measurement.

Regarding adaptive team-level phenomena, CAST measurement is more direct and less inferential than approaches based on knowledge elicitation (Cooke and Gorman 2006). Ideally, TSA measurement should take place when operators encounter novel or unlikely situational constraints beyond the common and valued state space of the DC2 environment (i.e. introspectively 'under-represented' states of the DC2 environment). The actions, comments, behaviours, and interactions of team members are then documented with respect to the TSA researcher's global perspective on both situational constraints and the DC2 environment. TSA is measured as the team's coordinated response to situational change, with unanticipated change providing the most sensitive test. So far we have identified five concepts which should be operationally defined in order to construct a CAST instrument:

1. **Identify roadblocks**: unlikely events requiring adaptive and timely team-level solutions.
2. **Document primary perceptions**: unless absent, asleep, or otherwise distracted, one or more team members in an ad hoc DC2 environment perceives some aspect of the roadblock and can react to it.
3. **Document secondary perceptions**: an operator who is attuned primarily to his/her own perspective on the roadblock has the ability to experience roadblocks in new ways by interacting with other operators who are more attuned to other aspects of the DC2 environment.
4. **Document coordinated perceptions**: this involves the reciprocal effort involved in team members explaining what they are perceiving to each other in order to put together a picture of the situation that is more than just the sum of their individual awareness.
5. **Document coordinated actions**: the actions of team members are often constrained by what other team members are doing, as well as by the situation itself; to some extent, operator – environment actions need to be coordinated in order to address specific roadblocks.

2. Case study: Operation Provide Comfort friendly fire incident

Following the 1991 Gulf War to oust the Iraqi invasion of Kuwait, coalition forces in Iraq launched Operation Provide Comfort (OPC), a 17-month-long campaign involving mostly US Air Force (USAF) sorties into the northern no-fly zone (NFZ). While OPC was essentially designed to act as a deterrent to the encroachment upon Kurdish lands in the north by Iraqi troops in the south, numerous humanitarian and peace-keeping missions were incorporated in order to counteract the plight of refugees in the northern NFZ. On 14 April 1994 a peace-keeping mission ended in tragedy as 26 people, including a multinational contingent of peace-keepers and the crews of the two US Army Black Hawk helicopters transporting them, were mistakenly shot down by two USAF F-15Cs performing a routine morning sweep of the NFZ. Although factors at multiple levels of socio-technical design have been implicated as contributing to this tragedy (Leveson *et al.* 2002, Snook 2002, Alexander *et al.* 2004), it is also important to identify causes stemming from DC2, in particular by measuring the TSA of the operators involved. We will address TSA in the ad hoc DC2 team consisting of the two F-15Cs and Airborne Warning and Control System (AWACS) controllers in charge of coordinating all flights to and from a NFZ tactical area of responsibility (TAOR) using concepts from both CAST and knowledge-based measurement. For a more complete account of OPC, and this accident in particular, we refer readers to Snook (2002).

Every morning during OPC USAF jets swept the TAOR, scanning for Iraqi ordnance. It was a well-known mandate that no coalition aircraft were to enter the TAOR before this sweep was completed. However, the two Black Hawks had obtained special clearance allowing them to enter the area during the morning sweep, a crucial fact left out of the briefings of the two F-15C pilots. Compounding this unusual situation, because of a computer malfunction of a radar console the AWACS en route controller, who normally sat next to the AWACS TAOR controller, took up station at a console which physically separated him from the local environment of the TAOR controller: '. . . it effectively removed him from physical contact with the TAOR controller. This . . . fact is particularly important, because the enroute controller physically "passes" control of all friendly aircraft to the TAOR controller as each flight enters the [NFZ]' (Snook 2002, p. 117). However, accepted practice did not dictate that, upon entering the TAOR, communication from the Black Hawks should be explicitly handed off from the enroute controller to the TAOR controller, who usually sat side by side and could glance at each other's screens (e.g. 'practical drift' within the system (Snook 2002)). Under usual circumstances this would have been a trivial non-event, as friendly aircraft have distinctive identify friend or foe (IFF) signatures and rarely travelled in the TAOR during the morning sweep. However, partly because of the computer malfunction and partly because of a lack of operator interaction, the TAOR controller, and ultimately the F-15Cs, performed as usual and were never aware that these aircraft were friendly Black Hawks.

As the two F-15Cs began their routine sweep of the area they detected the two helicopters on radar. After successive attempts to identify them electronically, as per procedure, they received confirmation that the TAOR controller was receiving 'unidentified' rather than 'friendly' IFF signatures from these aircraft. The two F-15Cs then initiated a visual scan, reporting the presence of two Soviet-made Iraqi Hind helicopters. This identification failure involved poorly illustrated visual identification materials, because the Black Hawks had an unusual configuration with wing-mounted fuel tanks, and was exacerbated by the fact that the F-15C pilots' mental model of the Iraqi Hind included wing-mounted fuel tanks, while their model of Black Hawks did not (Leveson *et al.* 2002). The two F-15Cs shot down the two Black Hawks, with the lead pilot taking out the rear helicopter and his wingman taking out the lead helicopter, resulting in the deaths of all 26 people aboard. Although there were factors at many levels contributing to this tragedy, it seems likely that this unfortunate event could have been avoided if the Black Hawks had been interacting with the same AWACS controller as the F-15C pilots (the TAOR), or if the enroute controller, who normally sat beside the TAOR controller, had explicitly informed the TAOR controller about the unusual presence of two friendly helicopters in the NFZ that morning. Even more troubling, if the state of the DC2 environment had been evolving within its usual constraints, this lack of explicit coordination was considered adaptive in terms of reduced operator workload (Leveson *et al.* 2002) via compliance with a SMM of the environment, i.e. via implicit coordination (Entin and Serfaty 1999, Stout *et al.* 1999).

3. CAST versus. knowledge-based measurement of team situation awareness

3.1. *Case study analysis*

Table 1 lists three roadblocks and the operators most affected by them. From a knowledge-based perspective, the ad hoc team consisting of the AWACS controllers and the F-15Cs independently perceived various aspects of the roadblocks, and independently

Table 1. Case study: aggregate knowledge-based and CAST team situation awareness.

Roadblock	Most affects	Individual operator level: Primarily perceived by[a]	Individual operator level: Knowledge-based SA	Team level: Coordinated perception[a]	Team level: Coordinated action[a]
Computer malfunction	Enroute TAOR	Enroute TAOR	Switch consoles[b] Status quo coordination[c]	Enroute TAOR	None
Fuel tanks on wings	F-15Cs	F-15Cs	Iraqi Hind[b] Identify unfriendly[c]	None	None
Pre-sweep clearance	TAOR F-15Cs	Enroute	Friendly aircraft[b] Shared display[c]	None	None

[a]CAST component; [b]heterogeneous SA requirements; [c]overlapping SA requirements.

constructed a situation model based on a sequence of activated schemas and mental models related to their knowledge of the DC2 environment. If we accept the notion of a schema or mental model underlying the development of TSA (Endsley and Jones 1997), then each operator's mental model was accurate given a normal state of affairs but inaccurate given the unusual state of affairs, i.e. their situation model was flawed. Pursuing this line of reasoning, if their situation model rests upon their mental model, then in retrospect their mental models must also have been flawed with respect to this unusual situation. This poses a question with respect to the utility of the mental model concept in assessing TSA in highly dynamic environments.

The F-15C pilots and the TAOR controller maintained status quo coordination corresponding to their shared introspective SA requirements of overlapping enroute–TAOR SA requirements and overlapping AWACS display, as well as the redundant attempts between the F-15Cs and AWACS to identify the helicopters as friend or foe (table 1). The performance of each of these tasks was accurate given the operators' mental models and SMM of the DC2 environment (i.e. with the system not in failure mode (Baxter and Bass 1998)). Limiting the analysis to the ad hoc DC2 F-15C–AWACS team, by summation across the team members' 'not in failure mode' knowledge of the system, a knowledge-based measure would point to adequate TSA in this case study. In fact, none of the operators mentioned here faced a court martial following this incident: 'Here it is "inaction" that needs explaining—the collective crimes of omission, rather than individual sins of commission' (Snook 2002, p. 103). Paraphrasing Snook, individually the operators did their jobs while exhibiting adequate heterogeneous and overlapping knowledge-based SA, but collectively they did not address the unusual situation. The senior AWACS director, who did face a court martial as a result of this incident, could have coordinated the situation, but was never properly a part of the team (the director was not held responsible (Snook 2002)). If this senior director had been responsible, we could have queried him and we might have been able to conclude something about a lack of TSA in terms of knowledge-based measurement. Properly, however, this would correspond to centralized command and control and fails to address TSA in network-centric DC2 environments.

A brief analogy (Gibson 1966) can be used to illustrate the difference between the CAST approach and the knowledge-based approach to measuring TSA in this case study. When we experience fire using our five senses, we see flames, smell smoke, feel the heat, taste ash, and hear the crackle. If we think of our five senses as heterogeneous DC2 operators, in order to perceive the global event 'fire' accurately, beyond the five independent sources of knowledge there must also be coordination across the senses. Similarly, when you are backing out of a parking garage it is convenient when your passenger notices that you are unwittingly encroaching on a large cement pole, but only to the extent that he/she communicates this fact to you. The operators in the OPC case study each perceived different aspects of the roadblocked environment, but for the most part these individual perceptions were not coordinated with each other (this is represented in the two rightmost columns of table 1).

In terms of team process behaviours, although the enroute controller demonstrated accurate individual SA concerning the Black Hawks entering the TAOR, he never tested his awareness with respect to coordinating his perspective on this information with that of the TAOR controller; he simply assumed that the TAOR controller saw the same thing via their overlapping display. In terms of CAST measurement, this is precisely the type of assumption that lends itself to poor TSA.

In short, the CAST approach proposes that TSA is not just the sum of operator SA. TSA is beyond the scope of adding up operators' private awareness of the situation and is predicated on operator interaction. Accordingly, we posit that good TSA does not depend on all operators being individually aware nor does it make sense for everybody to be aware of the same thing (in fact, this may be overkill for DC2 environments). We believe that what is most important is that the right information is communicated to the right team member at the right time, and this involves team coordination.

3.2. *Practical comparison*

Leveson *et al.* (2002) cited inaccurate mental models on the part of the F-15C pilots as well as the TAOR controller as contributing factors in this accident. It can be argued that this is consistent with a lack of TSA from both the CAST and knowledge-based approaches. However, a crucial difference lies in the causal mechanism for explaining poor TSA. CAST focuses on deficient team coordination, whereas knowledge-based measures focus on the inadequacy of a situation model based on individual operators' mental models. A knowledge-based theory might say that an accurate situation model *should* have been constructed for the unusual situation based on the primary perceptions and stored knowledge/mental models of the operators. The assumption that follows is that individual SA related to unusual environment states is *created* from the stored knowledge of operators, rather than *discovered* through operator interactions. However, in the OPC case study, a TSA researcher is left in the untenable position that TSA was *accurate* in terms of operator knowledge of environment states, as long as the operators did not have an accurate mental model of the unusual situation, but *inaccurate* if they had an accurate mental model of the unusual environment state. These conflicting SA requirements were beyond the responsibilities of the operators involved in the case study (Snook 2002), and this presents an ontological (i.e. 'chicken and egg') problem for the knowledge-based approach.

In practice, if we were to attribute a lack of TSA post hoc to an operator's mental representations (e.g. introspection) rather than maladaptive team coordination, we might be misled to conclude a lack of individual operator foresight instead of deficiencies in team coordination. Using Snook's (2002) terminology, we might be misled to conclude 'individual sins of commission' rather than 'collective crimes of omission'. As we have noted, the utility of DC2 systems rests in the capacity for adaptive emergent team coordination across a division of labour. In knowledge-based approaches, the reliance on the concept of a dynamic situation model constrained by a static mental model largely belies the concept of being 'adaptive' or 'emergent'. In addition, the notion of 'overlapping SA requirements' is conflated with the notion of a heterogeneous division of labour. In contrast, CAST measurement does not entail assumptions about aggregated knowledge-based awareness; rather, it assumes that TSA is predicated on seeing a roadblock in differentiated ways coordinated through operator interactions.

4. Conclusion

We have presented a theoretical basis for measuring TSA in DC2 environments by building on earlier notable discussions of SA measurement in the human factors/ ergonomics literature (Endsley 1988, 1995, 2000, Flach 1995, Smith and Hancock 1995, Sandom 2001) and by introducing some practical issues involved in measuring TSA in network-centric DC2 environments. Specifically, in DC2 systems heterogeneous team members are not expected to have a complete picture of the state of the DC2 environment at all times, especially under trying conditions, and this necessitates interaction between operators. In this regard, adaptation to any 'uncommon and unvalued' restructuring of the DC2 environment, through a self-governed change in team coordination, may contain a great deal of information concerning TSA. Aggregate knowledge-based measures of TSA do not tap this type of information because they do not directly tap team-member interactions. However, by using the basic components of CAST measurement, this information can be extracted by TSA researchers in order to gauge TSA in DC2 environments, allowing new issues to be addressed. For instance, recent team research implies that cross-training for a SMM or introducing shared displays may not always be effective in improving performance in heterogeneous teams (Cooke *et al.* 2003). Rather, the thesis that operators can dynamically modify each other's perceptual and action capabilities by interacting with each other, lends itself to an entirely different set of implications concerning team cognition, including implications for TSA theory, measurement, training, and design.

4.1. *Implications for TSA theory*

The CAST measure is inspired by ecological views of psychological phenomena, wherein unique perspectives on the same information are coordinated through interaction. Knowledge-based query measures are inspired by the classical information processing view of psychological phenomena, wherein internal models play the primary role. Therefore the focus of attention is on interaction processes in the former case, and 'in the cognition of the human mind' (Bolstad and Endsley 2003, p. 369) in the latter. Likewise, measurement focuses on the coordination of interactions in the former case, and individual and aggregate knowledge in the latter. However, thinking about TSA from these two perspectives has implications for how we think of team cognition more generally. The fact that we can observe team interactions more readily than individual or aggregate knowledge gives approaches like CAST the advantage of observing team cognition directly, rather than drawing inferences about it based on aggregated individual measures.

4.2. *Implications for TSA measurement*

It is difficult to elicit knowledge (including current understanding or projections of a situation) without interrupting the task and thus interfering with ongoing TSA-related processes. Therefore query-based metrics tend to be somewhat retrospective (after the processing has occurred) and introspective (the task is interrupted and often the effective stimulus is removed). Alternatively, because the CAST measure focuses on interactions, it is task based and can be taken *in vivo* while the task is in progress and without interruption (Gorman *et al.* 2005).

The information gleaned from queries and other knowledge-based measures is usually in the form of SA or TSA accuracy, and, in the case of TSA, sometimes intra-team similarity. A CAST measure can likewise assess accuracy or deviation from an optimal solution, but the output is not an answer to a query (or a set of team answers), but a sequence of specific actions and interactions. Whereas query output reflects the *outcome* of situation assessment (cf. the 'manifestation' but not the 'cause' (Baxter and Bass 1998)), as the queries can be thought of as tapping a fleeting situation model that has recently been generated and refined, CAST output reflects the *process* of assessing the situation itself. Therefore query measures are more outcome oriented than a CAST measure, which is more process oriented. This measurement difference has important ramifications for DC2 systems in that it is difficult to diagnose the processes underlying poor TSA using query measures. Alternatively, by focusing on patterns of interactions, CAST measurement provides diagnostic information about the adaptability of team-member connections.

4.3. *Implications for TSA training*

Query measures of TSA focus on building a common picture. This is a reasonable training goal, but again is perhaps an unrealistic expectation, especially in large DC2 environments. Additionally, because query measures are outcome oriented, it is not clear exactly how one would train to achieve the goal of a common picture, i.e. it is unclear by what process the goal of a common picture is achieved. Alternatively, CAST suggests that training should focus on coordinated interactions among team members. One way to do this is through simulation, in which roadblocks are presented to the team in order to train for adaptive team coordination. Such training techniques might involve training in simulated domains such as distributed team video-gaming, which requires coordinated perception and action on the part of the team.

Assessment is also relevant to training. Although query measures can be used to assess knowledge-based weaknesses in TSA, they offer little diagnostic guidance. For instance, it is possible that for large heterogeneous teams query measures would underestimate or fail to discriminate process-based aspects of TSA as pathogens underlying inaccurate awareness of the situation. On the other hand, CAST may be able to make finer discriminations on the basis of whether specific interactions did or did not occur. Thus CAST or a similar measurement approach can be used to assess TSA by providing diagnostic information about a team's specific TSA weaknesses.

4.4. *Implications for TSA design*

Query and other knowledge-based methods for measuring TSA tend to suggest that TSA can be improved by providing team members with a common picture of the situation. This is often achieved by giving every member of the team displays or shared displays which contain all the information that is relevant to the team as a whole. In general, the idea is that if team members A and B need to have a more similar understanding of the situation, then they need to share displays. However, this design principle breaks down in

DC2 environments as the size of the team increases and as team members have more specialized roles, where it may be prohibitive and counteractive, respectively, to give everyone mutual access to the same information.

Alternatively, CAST suggests that, to improve TSA, tools that facilitate adaptive and timely information sharing across the DC2 environment are desirable. Adaptive and timely information sharing does not mean that everybody has access to the same information at the same time, but means communicating the right information (and no more than this) to the right person at the right time. Design principles which facilitate open-source DC2 network architectures are suggested by CAST, such that network connections evolve based on operator adaptations to various situational constraints.

Acknowledgements

The authors acknowledge Roger Schvaneveldt, Olena Connor, Harry Pederson, Janie DeJoodie, Polemnia Amazeen, and Preston Kiekel for their input concerning ideas expressed in this paper. Portions of this research were supported by the US Air Force Office of Scientific Research (Grant No. FA9550-04-1-0234) and by the US Office of Naval Research (Grant No. N00014-03-1-0580).

References

ALEXANDER, R., HALL-MAY, M. and KELLY, T., 2004, Characterisation of systems of systems failures. In *Proceedings of the 22nd International System Safety Conference*, pp. 499–508 (Unionville, VA: System Safety Society).

ARTMAN, H., 1999, Situation awareness and co-operation within and between hierarchical units in dynamic decision making, *Ergonomics*, **42,** 1404–1417.

ARTMAN, H., 2000, Team situation assessment and information distribution, *Ergonomics*, **43,** 1111–1128.

BAXTER, G.D. and BASS, E.J., 1998, Human error revisited: some lessons for situation awareness. In *Proceedings of the 4th Symposium on Human Interaction with Complex Systems*, pp. 81–87 (Los Alamitos, CA: IEEE Computer Society Press).

BOLSTAD, C.A. and ENDSLEY, M.R., 1999, Shared mental models and shared displays: an empirical evaluation of team performance. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, pp. 213–217 (Santa Monica, CA: Human Factors and Ergonomics Society).

BOLSTAD, C.A. and ENDSLEY, M.R., 2003, Measuring shared and team situation awareness in the Army's future objective force. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, pp. 369–373 (Santa Monica, CA: Human Factors and Ergonomics Society).

CONNOR, O., COOKE, N.J., WEINGER, M.B. and SLAGLE, J., 2004, Using psychological scaling techniques to assess clinical expertise in anesthesiology. In *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, pp. 1746–1750 (Santa Monica, CA: Human Factors and Ergonomics Society).

COOKE, N.J., 1994, Varieties of knowledge elicitation techniques. *International Journal of Human–Computer Studies*, **11,** 29–68.

COOKE, N.J., 1999, Knowledge elicitation. In *Handbook of Applied Cognition*, F.T. Durso (Ed.), pp. 479–509 (Chichester: John Wiley).

COOKE, N.J. and GORMAN, J.C., 2006, Assessment of team cognition. In *International Encyclopedia of Ergonomics and Human Factors*, 2nd edn, W. Karwowski (Ed.) (London: Taylor & Francis).

COOKE, N.J. and SHOPE, S.M., 2004, Designing a synthetic task environment. In *Scaled Worlds: Development, Validation, and Application*, S.G. Schiflett, L.R. Elliott, E. Salas and M.D. Coovert (Eds.), pp. 263–278 (London: Ashgate).

COOKE, N.J. and SHOPE, S.M., 2005, Synthetic task environments for teams: CERTT's UAV-STE. In *Handbook on Human Factors and Ergonomics Methods*, N. Stanton, A. Hedge, K. Brookhuis, E. Salas and H.W. Hendrick (Eds.), pp. 46.1–46.6 (Boca Raton, FL: CRC Press).

COOKE, N.J., STOUT, R. and SALAS, E., 2001, A knowledge elicitation approach to the measurement of team situation awareness. In *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*, M. McNeese, E. Salas and M.R. Endsley (Eds.), pp. 114–139 (Santa Monica, CA: Human Factors and Ergonomics Society).

Cooke, N.J., Kiekel, P.A., Salas, E., Stout, R., Bowers, C. and Cannon-Bowers, J., 2003, Measuring team knowledge: a window to the cognitive underpinnings of team performance, *Group Dynamics*, **7,** 179–199.

Craik, K., 1943, *The Nature of Explanation* (Cambridge: Cambridge University Press).

Dekker, S.W.A., 2000, Crew situation awareness in high-tech settings: tactics for research into an ill-defined phenomenon. *Transportation Human Factors*, **2,** 49–62.

Durso, F.T. and Gronlund, S.D., 1999, Situation awareness. In *Handbook of Applied Cognition*, F.T. Durso, R.S. Nickerson, R.W. Schvaneveldt, S.T. Dumais, D.S. Lindsay and M.T.H. Chi (Eds.), pp. 283–314 (New York: John Wiley).

Endsley, M.R., 1988, Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society 32nd Annual Meeting*, pp. 97–102 (Santa Monica, CA: Human Factors and Ergonomics Society).

Endsley, M.R., 1995, Measurement of situation awareness in dynamic systems. *Human Factors*, **37,** 65–84.

Endsley, M.R., 2000, Direct measurement of situation awareness: validity and use of SAGAT. In *Situation Awareness Analysis and Measurement*, M.R. Endsley and D.J. Garland (Eds.), pp. 147–173 (Mahwah, NJ: Erlbaum).

Endsley, M.R. and Jones, W.M., 1997, Situation awareness, information dominance and information warfare. Technical Report 97-01, US Air Force Armstrong Laboratory.

Endsley, M.R., Hansman, R.J. and Farley, T.C., 1999, Shared situation awareness in the flight deck-ATC system. *Aerospace and Electronic Systems Magazine*, **14,** 25–30.

Entin, E.E. and Serfaty, D., 1999, Adaptive team coordination. *Human Factors*, **41,** 312–325.

Flach, J.M., 1995, Situation awareness: proceed with caution. *Human Factors*, **37,** 149–157.

Fowlkes, J.E., Salas, E., Baker, D.P., Cannon-Bowers, J.A. and Stout, R.J., 2000, The utility of event-based knowledge elicitation. *Human Factors*, **42,** 24–35.

Franz, G.F., 2004, Decentralized command and control of high-tech forces: aligning practice with doctrine at the operational level. Report, Naval War College, Newport, RI.

Gibson, J.J., 1998, Visually guided locomotion and visual orientation in animals, *Ecological Psychology*, **49,** 161–176.

Gibson, J.J., 1966, *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton-Mifflin).

Gorman, J.C., Cooke, N.J., Pedersen, H.K., Connor, O.O. and 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).

Hackman, J.R., 1987, The design of work teams. In *Handbook of Organizational Behavior*, J.W. Lorsch (Ed.), pp. 315–342 (Englewood Cliffs, NJ: Prentice Hall).

Hansen, D.K., 2004, Can decentralized command and control complement network-centric warfare? Report, Naval War College, Newport, RI.

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

Ilgen, D.R., Hollenbeck, J.R., Johnson, M. and Jundt, D., 2005, Teams in organizations: from input-process-output models to IMOI models. *Annual Review of Psychology*, **56,** 517–543.

Leveson, N.G., Allen, P.M. and Storey, M.A., 2002, The analysis of a friendly fire accident using a systems model of accidents. In *Proceedings of the 20th International System Safety Conference*, pp. 345–357 (Unionville, VA: System Safety Society).

Prince, C. and Salas, E., 2000, Team situation awareness, errors and crew resource management: research integration for training guidance. In *Situation Awareness Analysis and Measurement*, M.R. Endsley and D.J. Garland (Eds.), pp. 325–347 (Mahwah, NJ: Erlbaum).

Proctor, M.D., Panko, M. and 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,** 191–205.

Rasker, P.C., Post, W.M. and Schraagen, J.M.C., 2000, Effects of two types of intra-team feedback on developing a shared mental model in command and control teams, *Ergonomics*, **43,** 1167–1189.

Rentsch, J.R. and Klimoski, R.J., 2001, Why do 'great minds' think alike?: antecedents of team member schema agreement. *Journal of Organizational Behavior*, **22,** 107–120.

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

Sandom, C., 2001, Situation awareness. In *People in Control: Human Factors in Control Room Design*, J. Noyes and M. Bransby (Eds.), pp. 51–68 (London: Institution of Electrical Engineers).

SEAMSTER, T.L., REDDING, R.E. and KAEMPF, G.L., 1997, *Applied Cognitive Task Analysis in Aviation* (London: Ashgate).

SMITH, K. and HANCOCK, P.A., 1995, Situation awareness is adaptive, externally directed consciousness. *Human Factors*, **37,** 137–148.

SNOOK, S.A., 2002, *Friendly Fire: The Accidental Shootdown of U.S. Black Hawks Over Northern Iraq* (Princeton, NJ: Princeton University Press).

SOHN, Y.W. and DOANE, S.M., 2004, Memory processes of flight situation awareness: interactive roles of working memory capacity, long-term working memory and expertise. *Human Factors*, **46,** 461–475.

STEINER, I.D., 1972, *Group Process and Productivity* (New York: Academic Press).

STOUT, R.J., CANNON-BOWERS, J.A., SALAS, E. and MILANOVICH, D.M., 1999, Planning, shared mental models and coordinated performance: an empirical link is established. *Human Factors*, **41,** 61–71.

TAYLOR, R.M., 1990, Situational awareness rating technique (SART): the development of a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations*, AGARD-CP-478, Neuilly Sur Seine, France, pp. 3.1–3.17.

WICKENS, C.D., 1996, Situation awareness: impact of automation and display technology. In *Situation Awareness: Limitations and Enhancement in the Aviation Environment*, AGARD Conference Proceedings, Brussels, pp. k2.1–k2.13.