PAPER ID 126 -- 14th ICCRTS -- "C2 and Agility"

Assessments of IT's Support of C2

Topic 8: C2 Assessment Tools and Metrics

Name of Author(s): James Davidson[1], Alex Pogel[1], Jeffrey A. Smith[2]

Point of Contact: Alex Pogel, alex.pogel@psl.nmsu.edu

1: Name of Organization: Physical Science Laboratory, New Mexico State University

Complete Address:

Physical Science Laboratory,
MSC-PSL,
P.O. Box 30002,
Las Cruces, NM 88003-8002

Telephone: 575-646-9176
E-mail Addresses: jdavidson@psl.nmsu.edu, alex.pogel@psl.nmsu.edu

2: Name of Organization: U.S. Army Research Laboratory

Complete Address:

U.S. Army Research Laboratory
AMSRD-ARL-SL-EM
WSMR, NM 88002-5513

Telephone: 575-678-1332
E-mail Addresses: jeffrey.a.smith1@us.army.mil

Assessments of IT's Support of C2

Authors: James Davidson[1], Alex Pogel[1], Jeffrey A. Smith[2]

Organizations: 1: Physical Science Laboratory, New Mexico State University; 2: U.S. Army Research Laboratory, White Sands Missile Range, New Mexico.

Abstract:

The goal of information superiority presents an analytic challenge: how to assess information technologies' support to command and control. The assessment difficulty lies in the cognitive domain that the technology supports, and identifying those features of the cognitive domain that lend themselves to quantification and objective assessments. This paper describes solutions to this problem. First, we survey the Army's traditional human-in-the-loop IT assessment methods, and identify their limitations. Next, we discuss M&S' relative advantages in assessing IT support of C2. Our major argument emphasizes three essential requirements: definition, representation and objective measurement of key C2 products, specifically of situation awareness and situation understanding of tactical decision makers.

1. Introduction

Over the past decade the U.S. Army has diligently pursued an information superiority advantage for its Future Forces (2002; 2004) with a consistent focus on the information technologies (IT) that are expected to provide unprecedented levels of situation awareness (SA), a common operational picture (COP), strategic, operational and tactical agility, and ultimately improved force effectiveness across the spectrum of conflict. Consequently, IT spending constitutes a sizable portion of the U.S. Army's budget, for example the FY 09 Army Budget request included $5.3 billion for communications and electronics acquisitions and an additional $10.5 billion for research, development, test and evaluation, with a sizeable percentage earmarked for Future Combat System (FCS) IT enhancements (2009). Current IT-related line items portend major enhancements to the functions of Command and Control (C2) that were simply not possible just ten years ago. To highlight one example, Blue Force Tracker (BFT) provides current Brigade Combat Teams (BCTs) with a real time friendly force location capability. Other recent and future Army IT acquisitions include the Army Battle Command System, Joint Tactical Radio System, Force XXI Battle Command Brigade and Below, Tactical Internet, Unmanned Aircraft Systems and Unattended Ground Sensors, to name just a few.

Beyond enhancing contemporary *command* functions and expanding existing *control* capabilities, IT acquisitions have suggested shifts in the nature and form of C2 – a good example

(Alberts and Hayes, 2006). In a nutshell, the new sensing, communicating, and data processing capabilities have suggested an organization-wide version of Germany's *auftragstaktik* (Widder, 2002) that can be realized by coupling new IT acquisitions with the U.S. Army's preference for Mission Command (2003) and more broadly speaking, through the concept of Edge Organizations (Alberts and Hayes, 2003).

Acquisitions of the U.S. military typically include rigorous test and evaluation (T&E) processes, mandated and funded by Congress (U.S. Department of Defense, 2002), and this includes IT systems. There are at least three distinct types of IT assessments: those that test and evaluate 1) each component within a system for independent functionality; 2) the interoperability of various components within a system or of systems within a system of systems; and, 3) the degree to which IT supports command and control. There are many quantitative methods widely used for evaluating technology in the first two IT assessment types (Booz Allen & Hamilton, 1999; Darilek *et al*., 2001; Network Systems Group, 2009). However, technology is not an end to itself, and IT improvements are almost always oriented toward enhancing the performance of the consumers supported by the IT, most prominently, commanders and soldiers, so the third assessment type is the most essential. In other words, only the third assessment type directly involves the *raison d'être* of IT, namely its enhancement of C2. By implication, the more traditional component- and subsystem-focused DoD T&E process is insufficient and perhaps even inappropriate for the envisioned leader centric, network enabled system of systems operating environment (Charles and Turner, 2004; Lamartin, 2004; Starks and Flores, 2004). This view, that IT's core contribution is to enable commanders and soldiers, presents the central issue addressed in this paper: *how to objectively assess an information system's support to C2*.

The difficulty in producing assessments of IT's support to C2 stems from the dramatic difference in transparency witnessed between these two domains. In physical science and engineering developments in IT, even the most futuristic IT, there are independently verifiable criteria and corresponding measurement methods to support assessments. The cognitive domain, however, is understood through an array of micro-theories, with no single unified theory of cognition that establishes what, how and why things work within the minds of the warfighters who consume IT content (Newell, 1990; Newell and Simon, 1972). Exacerbating the situation, and related to the lack of a unified theory of cognition, is that there are no holistic measurement methods – not for the commander, leader, soldier, nor intelligence staff officer – that quantify sub-processes' contributions to the summation of each person's cognitive processes. We do not intend to present such a unified theory, rather our aim is to present some new ideas on how to assess IT's support of C2. In short, given the massive domain ontologies related to IT and C2 for a given warfighting context, especially at the level of a BCT with roughly 5000 separate IT content consumers, we believe the fundamental issues are

1. identifying the *ideal phenomena* to measure,

2. identifying the *best context* in which to measure the phenomena, and
3. determining how to *objectively measure* the phenomena.

In this paper we argue that a decision makers' Situation Awareness (SA) and Situation Understanding (SU) are the ideal phenomena to measure, as they represent the realization of the universally accepted tactical factors Mission, Enemy, Terrain, Troops Available, Time and Civil Consideration (METT-TC) in the decision maker's cognition. Further, we believe that SA and SU products are best measured within a well-constrained warfighting scenario, and that these products *can be* objectively measured if warfighting scenarios are represented via M&S technologies. The further sections of the paper present these arguments.

In Section 2, we review the Observer/Controller team paradigm, the U.S. Army's preferred method for assessing IT's support of C2, during human-in-the-loop (HITL) experiments at the U.S. Army's Combat Training Centers. In Section 3, we propose modeling and simulation (M&S) as an alternative assessment method that alleviates issues with HITL-based assessment methods, such as lack of transparency, subjective assessment methods and prohibitive costs. Sections 4, 5 and 6 are respectively devoted to *defining, representing and measuring* the fundamental and quantifiable cognitive concepts present within the C2 domain that are essential for the goal of measuring IT's support of C2. We emphasize the criticality of identifying those parts of the C2, IT and warfighting ontology that are to be represented, of committing to the level of fidelity to be achieved in the representation, and of objectively measuring C2 products.

2. Human-in-the-loop Assessment of IT's Support to C2

In this section we discuss the Observer/Controller (O/C) team paradigm employed during HITL experiments at the Army's Combat Training Centers (CTCs), as the primary method for assessing IT's support to C2. Since the mid-1980s the Army has used highly skilled warfighting experts organized as O/C teams at its CTCs to provide unit commanders with unit level operational readiness assessments. O/C teams are a collection of specially selected, seasoned warfighting professionals with subject matter expertise in all of the warfighting functions. These O/C teams are staffed to provide complete coverage of every Brigade Combat Team (BCT) key leader, commander and staff member in every warfighting function. In fact, the different O/C teams (one for each type of specialized battalion assigned to or in support of the BCT) are organized to mirror the structure of each organization they assess. Frequently the Army's senior leadership also directs the O/C teams to observe and assess specific warfighting trends that need improvement. The ineffective application of fires is one example:

> "The most common and significant trend O/Cs have observed in transitioning units (from Stability Operations after OIF or OEF to the full spectrum of conflict experienced at the CTCs) is that all cannon sections need more practice in executing their fundamental individual and section-level/collective tasks—from

> the fire support team (FIST)/combat observation lasing team (COLT) to battery/platoon fire direction centers (FDCs) to communications and maintenance of individual howitzer sections. For example FDCs that don't have an established crew drill can't troubleshoot a routine database error or process a digital fire mission." (Waters, 2004)

The O/C teams possess a vast array of data collection methods including audio and video transcription, enabling the O/Cs to sustain near ground truth SA with respect to the assessed BCT by using their own perfect communications system to exchange data with one another and with their counterparts in the Opposing Forces. To complete their common operational picture and to aid their after-action-review (AAR) preparation, the O/Cs can also call on the CTC's Tactical Analysis Facility (TAF). The TAF includes a data capture and visualization capability, a forerunner to, and inspiration for, the FBCB2 and Blue Force Tracker technologies that exist to provide each O/C with the exact position location and identification of every friendly, enemy and non-combatant force in the operational environment. O/Cs will often direct their TAF analysts to audit and store visual and auditory battle action summaries of decisive engagements, events or activities for subsequent use during AARs.

Other key systems within the CTCs that make it possible for the senior O/Cs (Brigade and Battalion) to perform their work are the Battle Damage Assessments (BDA) recorded by a supporting cast of TAF analysts who call upon subordinate company and platoon O/Cs, and the OPFOR to provide the BDA (and other combat) data essential to each senior O/Cs AAR. For example, O/Cs can collect the exact number of enemy tanks destroyed, by whom, when and where. Further, O/Cs can access data on rounds fired by munition type, gallons of fuel consumed, personnel killed or wounded, messages sent or received by particular nodes in a network, number of ineffective indirect fire missions, or any other combat relevant data deemed important for the lessons learned that senior O/Cs expect to extract during AARs. In a C2-specific example, it is possible for a given O/C (e.g. the Senior Brigade Staff Trainer) to describe who knew what (e.g. the source and description of a priority intelligence requirement), when, and whom the information was passed too, and when and how the information triggered a particular decision or action (or why the information should have led to a particular decision or action but did not). In short, O/Cs can access, manipulate and manage a vast amount of data and present their findings using a variety of visual techniques.

During their unit assessments, OCs shadow the commanders, leaders, staffs and soldiers of the BCTs training at the CTCs, then regroup and incorporate their local assessments into a system-wide unit appraisal. For units being assessed, the hallmark of the CTC experience is the unit After Action Review (AAR). The AAR is a well-crafted interactive process whereby the O/Cs draw out lessons learned from unit members through a series of probes backed by data. The O/C AAR succeeds when the members of a unit perform their own knowledge discovery. The O/C teams construct and orchestrate these AARs within hours of force on force or live fire missions and the AARs almost always involve the key leaders and staff of the BCT down to platoon level.

The AARs provide the BCT with immediate feedback in the form of casual chains while recently concluded missions remain fresh in the minds of participants. In the end, each AAR ultimately results in a number of key corporate and individual lessons learned that the BCT hopes to improve during its next battle with the OPFOR. In the next section, we chronicle some shortcomings of the O/C Team method, and suggest the use of Modelling and Simulation technologies as an alternative method that alleviates some of the fundamental problems.

3. The Use of Modelling and Simulation in IT Assessments

As one can imagine, the O/C assessment paradigm is extremely important to the Army. Significantly the CTC O/C teams are credited with helping many BCT's prepare for combat success, whether stability operations (e.g. OIF or OEF) or full scale war (e.g. Desert Storm). Yet the O/C team paradigm does present challenges to *objective* assessments of IT's contribution to C2. An undeniable fact is that O/Cs have human limitations in their capacity to collect, process, store, interpret and recall data in their desire to understand and then explain to a player unit what happened and why. It is also worth noting that O/Cs are prone to many of the same environmental stressors, e.g. fatigue, that player units experience since they keep the same operational tempo as the players themselves. In simple terms, O/Cs can never know ground truth. In addition, despite the fact that O/Cs shadow their counterparts, it is impossible for them to know precisely what and how their counterparts were thinking and when, since they are not able to read their counterparts' minds. Further, because the AAR is a post hoc knowledge discovery process, the participants are not likely to have total recall of their thought processes (not to mention the inaccessibility of their unconscious thought processes), nor the information available to them at the time of their decisions and actions. Another complication stems from the very public nature of the AAR wherein the entire chain of command is present. This situation inclines participants to put the best possible spin on problems in mission execution that are attributed to them. Also, as a unit's members execute successive missions, the player's will have learned something of the system and environment in which they function, thus denying repeatability. The presence of these and other factors naturally inclines O/Cs toward a rather qualitative and subjective assessment. Last, but certainly not least, is the cost of each CTC rotation, estimated to be in the range of $2-3 million for each month long exercise, beginning from the unit's deployment to the CTC and concluding upon its return to home station.

Two other assessment methods focused on extracting C2 measures from HITL experiments are the Situation Awareness Global Assessment Technique (SAGAT) interview technique and the System Knowledge score used by the Training and Doctrine Command of the U.S. Army. SAGAT (Endsley, 1988) is an interview technique used by cognitive task analysts (Crandall, Klein, and Hoffman, 2006) to assess a warfighter's level of situation awareness (SA) achieved during a HITL experiment; the levels of SA are often described as perception, comprehension and projection and referred to as Endsley's SA Levels 1, 2, and 3. The System Knowledge assessment score (Blechinger *et al*., 2004) is another measure intended to measure the timeliness, accuracy, and completeness of incoming information. These methods are designed for HITL experiments, and offer limitations similar to those seen in the O/C Teams, regarding transparency, repeatability and subjectivity. In summary, the primary pitfalls in producing assessments of IT's support of C2 via the O/C Team paradigm, or other experimental methods, are the lack of transparency of the humans being assessed, the subjectivity and cognitive limitations of the assessors themselves, the exorbitant cost of a unit's execution of a few missions, and the lack of repeatability.

We believe that it is possible and practical to perform *objective* assessments of IT's support of C2 in the comparatively inexpensive M&S environment with full transparency regarding the C2 entities modeled and without losing the ability to repeat experiments. Under the assumption that C2 entities can be sufficiently modeled via M&S technologies, the resulting analytic tool would have obvious benefits in terms of cost, transparency and repeatability, as well as offering unparalleled "what if" capability along with substantial time savings. However, there are in fact tremendous difficulties in modeling C2 processes absent from the human-in-the-loop case. Metaphorically speaking, HITL assessments of IT's support to real people exercising C2 is equivalent to taking a black-box viewpoint regarding the C2 processes in that analysts need not concern themselves with fidelity issues, since the IT, the content, and the consumers who exercise C2 based upon that content are all real. Further, the model of interactions between IT and consumers (and their human enemies), or the interaction between multiple consumers within a hierarchy, or the interaction between component IT systems is also not in question. Gaining access to, and discovering truths about, all these black-box interactions are serious difficulties, but the realism of the C2 processes, and most of the related phenomena, is not. Another M&S issue related to the previous point, but absent from the HITL case, is the necessity to expend considerable resources in defining the ontologies, choosing the relevant parts and specifying the requirements for the various systems and processes that are prevalent in the warfighting context under investigation. In HITL experiments, analysts simply see what is before them and rest assured that what they observe is as real as it gets. Nevertheless, in the overall balance of upsides and downsides, M&S is still a very attractive alternative to established HITL methods because of the transparent, repeatable and objective nature of the assessments that are possible. The next two sections present our thoughts on the problem of modeling C2 entities' decision processes, through careful definition and representation, in any M&S tool used to assess IT's support of C2, along with observations regarding the corresponding IT and environmental and domain concepts that must also be present.

4. Definition of Key Concepts in the C2 Domain

The first big challenge for C2 modeling concerns the term *command and control* itself. Like the term *system of systems*, much has been written about C2 and there are many definitions and points of view regarding C2; for example, the *Coalition Command and Control Bibliography* (Command and Control Research Program, 2009) lists well over 100 references as recommended reading on the topic of C2. Although there may be a lack of consensus about the definition and characterization of this large and highly complex fragment (i.e. C2) of a near infinite warfighting ontology, this does not prevent us from distilling the C2 essentials that lend themselves to quantification in M&S.

Taking the second term first, some have viewed *control* as something very mechanical and structured; control is often formalized as checklists and procedures that are usually action oriented, e.g. the steps taken in docking a cruise ship, landing a plane or running a nuclear power station. Others take an organizational perspective in viewing control as synonymous with supervising and directing people by observing behavior, acquiring feedback and making course corrections through a series of inputs to keep the organization moving toward stated objectives (note this view of control asserts indirect influence in contrast to the direct control by the pilot of an aircraft). The Army itself describes control as the processes, procedures, IT and facilities used (and in what ways) to support command. This point of view takes a systems perspective and includes concepts like procedures. Yet the real distinction to be made, in contrast to the mechanical or engineering perspectives is the centrality of the commander and the existence of control capabilities that enable the commander to exercise both direct and indirect influence over his unit, i.e. to command his unit, via a hierarchy of subordinates and staff in order to achieve mission.

With respect to the first term in C2, *command* is viewed as something less tangible, artful and personality dependent, relying on imagination, creativity, experience, intuition and social skills like motivating and morale building. Business might view command as something CEOs do, or might substitute *lead* and *manage* for command because the term command sounds militaristic to some. Technically the Army defines command as the legal authority and responsibility vested in a commander that the commander exercises, and that subordinates voluntarily recognize and accept, in guiding his forces toward mission accomplishment. This characterization is not very useful in M&S since it describes command as something a commander has, not something a commander does. Perhaps a more useful description (Alberts and Hayes, 2003) is a view of ***command*** (emphasis added) and control as an enterprise-wide example of establishing intent, determining roles, assigning responsibilities and relationships, establishing rules and constraints and monitoring and assessing the situation and progress. We also add allocating resources and making adjustments to any of the previous concepts deemed appropriate to achieve mission. Upon closer examination we find that this definition of command and control is completely consistent with the Army's definition of Battle Command, a term we paraphrase as the process of receiving and analyzing mission, assessing the current state, and then visualizing how to move the organization from its current state to the desired future state. Implied is the requirement to develop and communicate a concept (intent, objectives, roles, allocating resources, determining constraints, limitations, rules of engagement, etc.) that when properly executed is expected to achieve mission while continuously monitoring the situation and adapting the concept or any of its pieces, when necessary.

In looking at the various C2 definitions and the contexts in which they appear, *command* fundamentally involves decision making of various sorts, for example:

1. who to promote
2. who to discipline
3. who to motivate
4. what limited resources to allocate to whom and for what
5. what constraints to impose
6. what tactics to employ to accomplish mission

On the other hand, *control* fundamentally involves the use of IT and other tools to support and enable the functions of *command,* for example:

1. reconnaissance activities to collect pre-battle intelligence to support concept development,
2. data collection to support decisions regarding redistribution of limited fuel for an impending long distance advance,
3. friendly force position location information, relative to known enemy forces to support a commander's projection of near term futures,
4. sending the BCT operation's officer to provide an eyewitness account of a supporting effort battalion's breech of a complex obstacle belt,
5. Re-tasking a Unmanned Aerial System to overfly a major river to quickly find a suitable crossing site for friendly forces held-up on the near side.

In short, when exercising command by employing control, a commander makes enumerable decisions from the simple to the most complex, within both small and large contexts and over many decision domains e.g. personnel, technology, or tactics. Given the massive nature of the C2 domain, and the observation that all C2 activity is to enable tactical decision making (TDM), we deliberately focus our attention on tactical decisions and the critical IT content that supports the decision maker's cognitive processes.

One last point regarding the C2 domain, before we shift focus to tactical decision making, is the distinction between C2 considered as a collection of functions and/or capabilities supporting TDM, and C2 as an organizational strategy. The Army's preference for C2 is Mission Command (MC), a C2 strategy based upon subordinates exercising disciplined initiative within the framework of the commander's intent. In Mission Command, subordinates' adaptation is expected and their agility is enabled by widely distributing information and decision making authority throughout the organization; this is not too dissimilar from the ideas expressed in Edge Organizations (Alberts and Hayes, 2003). In contrast, Detailed Command is an alternative Army C2 strategy, considered appropriate in highly stable and easily controlled situations. Detailed

Command is based on a commander's specific decomposition and assignment of tactical mission tasks and his detailed control of their execution.

The Army's preference for Mission Command stems from a 1911 observation by von Moltke (Delbrück, as translated by Renfroe, 1980, p. 425) that "the advantage which a commander thinks he can attain through continued personal intervention [Detailed Command] is largely illusory". This fact is exacerbated in epistemically uncertain environments; in Detailed Command, the commander takes on an inordinate responsibility regarding task selection, execution and force agility, that is better shared with subordinates. Leadership is important in both C2 strategies and its importance in effectively applying combat power arises from the natural tension between a subordinates' desire for honor among his peers and self-preservation, and the commander's desire to accomplish his mission through the activities of his subordinates. In short, leadership aims to direct and supervise the execution of a concept of operation designed to achieve mission by influencing subordinates to perform in ways they may not otherwise prefer, if left to their own devices. In Detailed Command, the burden is squarely on the commander to overcome this tension, whereas in Mission Command, subordinates retain considerable flexibility in resolving the tension themselves, which promotes agility of the entire organization. The key point is that the continuing presence of these competing theories (and their hybrids) of C2 strategy indicates that commitment to one or another of these organization-level C2 strategies is *not* a universal feature. This tangential discussion barely touches on all there is to say about the Mission Command C2 strategy, that is itself just a fragment of the overall warfighting landscape but is also quite elaborate in its own right: for example, the Army describes Mission C2 in Field Manual 6-0 (2003); a publication with 6 chapters, 6 appendices, and 294 pages of text, not including sources and a glossary, and with over 50 illustrations. As mentioned above, we focus now on tactical decision making.

At a high level of abstraction the tactical decision making process (TDMP) begins with receipt of mission. The mission statement itself will contain who, what, when, where and why (the "5 Ws"). The first step in TDM is to perform a thorough mission analysis to identify the specified and implied tasks and to extract the essential tasks from the list that when completed would lead to mission accomplishment. The term METT-TC (short for **M**ission, **E**nemy, **T**errain, **T**ime - **T**roops available, **C**ivil considerations) helps decision makers to remember the factors they must consider in performing their mission analysis. Enemy follows Mission and for good reason because the enemy always has a vote and as the dictum goes, no plan is likely to survive the first meeting with the enemy. Notice that awareness of the enemy's location and strength and disposition and type are universal requirements in the TDMP. The next factor, after enemy is Terrain. Here, too, awareness of terrain features in terms of observation, key terrain (that which gives the controller a tactical advantage), mobility corridors, trafficability, soil types, weather conditions, line-of-sight, avenues of approach, light conditions, hydrology, etc. are also universal factors that do not depend on personality of the decision maker, form of C2 or type of conflict. Time follows terrain and refers to the planning and execution time available and the time until the mission needs to be completed. Next is Troops Available, which concerns itself with very same issues enumerated under Enemy above, but instead from the friendly vantage point, e.g. number and type, morale, percent strength, location, disposition, etc. Notice again that awareness of Troops Available factors do not depend on form of C2, personality of commander or type of warfare. Finally Civil Considerations, which is the awareness of all governmental, non-governmental and non-combatant issues of concern that could impact mission planning, execution and accomplishment. Once again, awareness of these elements is independent of personality, C2 strategy or form of warfare, albeit Civil Considerations are more pronounced in stability operations.

A warfighter's TDMP can be lengthy, analytic and deliberate when time is available and, to an outside observer, the TDMP frequently reflects the scientific problem solving process or rational decision making process. In the deliberate, prescriptive case, a commander and/or his staff (in recognition of the commander's intent and planning guidance) would develop several courses of action and then wargame and evaluate each one. Criteria are then developed for use in comparing each COA in order to determine the best choice. Once chosen, orders are produced and communicated, rehearsals performed and adjustments made; included in the plan are contingencies that answer the "what-if" analysis from wargaming.

From our previous C2 and TDM summary, we observe that the Army's longstanding decision-making factors Mission, Enemy, Terrain, Time, Troops Available and Civil Considerations (METT-TC) are universals of the C2 domain, as relevant to tactical decision making in the future as they are today. Situation awareness (SA) is a decision maker's available data drawn from the information system, while situation understanding (SU) is the decision maker's comprehension and projection of situation in light of his SA and Mission. For example, an SA product is location and type of friendly and enemy forces at a particular moment, while SU is the result of using that data to interpret mission and survivability, for example, "if I do not move quickly, using cover, to block the enemy, then he will interfere with the main effort I am to protect." These SA and SU products represent the realization of a decision maker's knowledge base and understanding, in a particular context, of the general METT-TC factors. Consider the emphasis on METT-TC factors in this ageless military dictum:

> "Therefore I say: 'Know the enemy and know yourself; in a hundred battles you will never be in peril. When you are ignorant of the enemy but know yourself, your chances of winning or losing are equal. If ignorant both of your enemy and of yourself, you are certain in every battle to be in peril'." (Sun Tzu, as translated by Griffith, 1963, p. 85 in the paperback version)

By contrast, the following passage describes specific SA and SU products realized as part of a specific warfighting context.

> “Sweeny patiently correlated the information that Dagger and CIA teams were sending back from eastern Afghanistan with technical information on the location and movement of Al Qaida forces provided by spy planes and satellites. Sitting in Dagger’s secure, compartmented intelligence facility, or SCIF—a long tent, surrounded by razor wire, guarded by 10th Mountain troops and permeated by the smell of burnt coffee—Sweeney worked to paint a picture of Al Qaida’s network of safe houses, transportation nodes and escape routes out…” (Naylor, 2005, p. 25)

Across the large body of literature on C2, ranging over millennia, from Sun Tzu to Clausewitz and beyond, the METT-TC factors and their realizations, SA and SU, are common ingredients in all C2 and TDM ontologies, independent of any particular commander, warfighting context or theoretical view of C2. However, abstract concepts like the METT-TC factors cannot be measured, while concrete data points and propositions about an environment (SA and SU) can be measured in terms of accuracy, timeliness, relevance, and uncertainty. In battle, the role of IT is to support SA and SU and hence TDM. Specifically, we assert that assessments of IT’s support to SA and SU should concern themselves with the perceived facts of the matter, independent of widely varying aspects of C2, such as leaders' personalities, culture, C2 strategy and nature of warfare. There is little lasting merit in analysts making judgments about IT’s support dependent upon to a decision maker’s experience, creativity or intuition. For example if a commander decided to deceive an enemy force by faking an attack to divert the enemy’s attention from the real, main attack, (assume that the scheme worked brilliantly) we would not look to IT to ask how IT supported this creative act inside the head of the commander. On the other hand, we would rightfully assess IT’s support to the commander by asking what percent of enemy forces where actually known by the commander and their location. Therefore, SA and SU are those cognitive products universally found in C2 activities that can, and should, be modeled and measured to provide IT assessments.

Combining this argument with the advantages of M&S described in Section 3, we conclude that the quantification and objective measurement in M&S of SA and SU are requirements for assessing IT’s support of C2. In the next section, we turn our attention to *representation* of these C2 concepts and then describe our *measurement* of these concepts in M&S as a basis for IT’s support of C2.

5. Representation of Concepts Central to C2 and IT

To take up the problem of using M&S to assess IT support of C2, one must first grapple with the assumptions imposed upon the M&S development, since modelling decisions made in light of these assumptions ultimately determine the type of analyses the M&S product can enable. Such grappling is witnessed, for example, in the analytic assumptions and limitations determined by

Hodges (1994), in an analysis of a series of training exercises conducted at the NTC, focusing on various reconnaissance elements in several units. To enable a full range of experiments devoted to IT assessments, we need an M&S tool that allows IT variations, whether due to intentional variation of technological alternatives or perturbations such as computer network operations or natural failures, to be assessed through measurements of the C2 products the variations produce. Our challenge is how to create such a model that supports not only the technological side of the Dynamic Model of Situated Cognition (Miller and Shattuck, 2004), that describes processes executed with the radios, wireless protocols, routing, etc., but also the cognitive side, including awareness, planning, deliberation, decision making and responses. Indeed, beginning with minimum requirements, it is clear that we must model carefully chosen aspects of C2 entities' products, and we must model the IT processes that support their production.

Traditionally, to represent decision makers in the M&S context, analysts relied on decision tables that in essence imposed an external decision making structure on the agents in a simulation (Hewett and Leuchner, 2003). Analysts go to great pains to create these tables by interacting with subject matter experts in defined, well-structured environments to elicit the underlying reasons for significant choices and then codify these responses in an "approved" solution. Yet, this process is difficult even when models of a domain are simple (Russell *et al.*, 2006). While eminently useful for questions involving different force packages, since constrained decision making reduces variation in outcomes and keeps the spotlight on the force variations being assessed, we must ask ourselves: is this externally imposed decision making structure sufficient for assessing the consequences of variations to IT performance when these consequences will be measured in the cognitive domain? Furthermore, can we convince ourselves that it is possible for any analyst to create a set of decision tables that captures all the manifold interactions in an environment and that also includes an information system subject to perturbations, such as device failures or adversary attacks? Some studies indicate the answer to this question is no, e.g. Jensen *et al*. argue that "[u]nder a broad range of circumstances, these approaches exhibit a nearly linear relationship between training set size and tree size, even after accuracy has ceased to increase." (1997)

We assert that the general answer to these questions is "no", based on our view that tactical decision making is fundamentally a dynamic process, something that cannot be *a priori* determined; because of the dynamic aspect, any modeling and simulation activity of tactical decision making must explicitly consider information, not as an abstraction but as meaningful, content-laden data, interpreted with respect to the current goals of the commander being modeled. Thus an M&S tool that supports credible analysis of IT's support of C2 must model the creation, transmission, receipt and understanding of information, so that the consequences of IT variations can be properly assessed. The model itself must generate the information, and the model must include processes that represent the interpretation of that information. Thus, scouts in a simulation of combat can no longer send a pre-defined message to their troop leader; they must send a situation dependent spot report: "I have three enemy tanks at grid location … Over."

Yet, this is not the whole story; the troop leader receiving the spot report must recognize the content of the report and realize that its relation to an important priority intelligence requirement specified by the battalion commander. By implication, this means that a spot report from a scout, sent up the chain of command could cause the battalion commander to shift his axis of advance down the left branch of a road instead of the right owing to the perceived threat.

In fact, models can and must generate this kind of information in order to assess IT's support of C2. Simulated decision makers must recognize and consume that information locally, that is, they cannot rely on an externally imposed decision structure to decide *in absentia*. Modelled decision makers must carry within themselves not only the means to evaluate the information, in terms of its relation to mission and self-preservation expectations, but also a means to decide some course of action once their evaluation is complete. Each such decision leads to actions in the simulated environment, that leads to more sensing, more communicating, more delivery of METT-TC content, and further decisions; an M&S environment implementing such a decision-action-perception-decision cycle offers a rich environment for experimentation related to IT's support of C2. Note, however, that we have not insisted on any particular level of fidelity for maneuver, engagement and battle damage models; indeed, to meet the goal of assessing IT's support of C2, these models need only exist in a form sufficient to support the continuation of the decision-action-perception-decision cycle.

As we remarked above, there is obviously too much phenomena in the C2 domain for any M&S tool to manage, and the requirement that modeled agents recognize, interpret and ultimately use information opens the question: how much of the decision making domain must one model? In the previous section, we have already isolated SA and SU products as fundamental enablers of all C2 activity; by contrast, we also identified C2 strategy and identification of tactical choices as too personal, too variable, to use as the measured criteria on which to base an assessment. In order to facilitate creation of the decision-action-perception-decision cycle, we require that simulated decision makers do make tactical choices, do conduct their operation under one or another C2 strategy, but these modelled phenomena should not be the focus of measurements, instead SA and SU products provide the foundation of assessments of IT’s support of C2.

Many previous research papers indicate that decision variables and decision outcome alternatives are limited. Graham *et al*. (2007) cites the work of Gigerenzer *et al*., (1999) to assert that human decision makers process only a few alternatives when facing a decision. They further cite Wickens *et al*., (2004) to assert that humans typically consider three or four attributes when making decisions, and that often, experts reduce the number of variables they consider to three or four categorical variables (Klein *et al*., 2007; Wiener and Nagel, 1988). Finally, they assert that there is little to be gained from decision making models with more than 10 decision making variables. The implications are that the modeller need not model all possible combinations of inputs and outcomes, but only model those foundational C2 products, namely SA and SU, that serve as inviolable bridges between the near infinite possible inputs and outputs. The research works mentioned above suggest that low dimensional models of these SA and SU products are sufficient to represent the corresponding products actually produced by experts. Further, since the use of IT is to reduce the uncertainty present in SA and SU products, these low dimensional models of information and information processing must incorporate uncertainty. The presence of explicit models of uncertainty provides the proper basis for decision making models to dynamically apply IT assets toward the reduction of that uncertainty.

Consequently, as we restrict our decision making models of a tactical decision maker, we must at least include a large number of the METT-TC factors identified in the previous section to establish representations of SA and SU that we have concluded are fundamental for assessments of IT's support of C2. One can create models of decision making, rooted in METT-TC data mentioned above, that can be used to assess perturbations to information systems (Bernstein, Jr., Flores, and Starks, 2006). These models by necessity must gather the primitives available to them as atoms of situational awareness. They must use this situational awareness data and an interpretation of that data, in light of their mission statement, to form an understanding of their current local situation, and then they must hypothesize futures from this local situation, ultimately deciding whether or not to remain on the current course of action or opting for a new course of action. While it is clearly difficult to model such a system, Kunde (2005) and Kunde and Darken (2005; 2006) suggest that models that explicitly incorporate the ability to perceive, hypothesize, and to decide based on information are more robust.

However, our discussion of M&S requirements is not yet complete; we must also combine these METT-TC factors with models of IT processes and corresponding methods that allow a decision maker to recognize variations in IT performance and to execute appropriate responses. We suggest that models of IT that include performance variation have two components, the performance variation itself, and an attempt by the system to restore to some degree of the *status quo ante*. For example, if jamming is observed on one frequency, a typical response of an entity in a C2 network attempting to use that frequency to communicate is to change frequencies, thus possibly mitigating the jammer. This is an example of a decision that must be placed in the hands of the simulated agent and not controlled by the analyst outside the simulation. As one might expect, it is not possible to model every such perturbation-response pair, so the modeler's choice, in conjunction with the analyst and in light of the specific question being addressed, is to model those perturbation-response pairs that are sufficient to contribute meaningful insight to the answer of the question. In rule form, two such perturbation-response methods might be:

1. In the presence of communications jamming, change frequencies, and if time permits, identify the source of the jamming and engage if possible
2. Having identified a network intrusion, identify the source of the network perturbation, and initiate appropriate counter response.

These choices imply, for example, that our model of a radio that supports communications must also be capable of modeling a "change in frequencies" or that the network can support intrusion detection and mitigation strategies. Effectively establishing these modeling choices is a matter of identifying the central functionalities of the technology of interest and entities' use of those functionalities, including mitigations of failure modes. The next section presents our ideas on how to use the transparency of C2 entities' decision processes, when modeled via M&S methods, to produce objective assessments of IT's support of C2, using an extension of the O/C Team paradigm employed at the CTCs.

6. Objective Measurement of Modelled C2 Processes

Having established the importance of focusing on the SA and SU products, supported by IT, that enable TDM, we turn our attention to the problem of producing *objective* measurements of these products. First, we note that an appealing and immediate response to the problem of IT assessments is for an analyst or subject matter expert to judge the quality of the most advanced of a C2 entity's products. The analyst judges these advanced C2 products, e.g. tactics, control, leadership, by employing his own temperament, experience or expertise; a natural tendency is for the analyst to judge IT favorably provided that its usage leads to decision activity the analyst prefers. The difficulty here is that these qualitative judgments are fundamentally personal in nature, and in unit level assessments that involve multiple decision makers and assessors there is often broad disagreement between the assessors themselves regarding the best tactics to employ in a particular situation. Since there are no textbook solutions to military decision problems, absolute judgment is simply not an objective assessment methodology.

In this section, we present a method that provides objective assessments of C2 products, that we call the Objective Information System Assessment method (OISA). In this method, we relativize analytic judgments regarding C2 products such as SA and SU, and even tactics, thereby avoiding various difficulties with the subjectivity of absolute judgments. The idea is to create comparisons between the subject decision maker, employing a particular suite of IT capabilities, and *what that same subject decision maker would have produced if offered an alternative data stream*. By construction, this self-referential assessment of the subject is based on the same view of mission, the same (potential) interpretation of raw information, and the same training, experience, and temperament. The greatest benefit of this approach stems from the objectivity of the assessment: any single battle will be fought by and assessed from the point of view of the decision makers involved, not by detached observers who, trusting in their own experience and temperament, believe they know better what needs to be done, if only given command of the assessed unit. We can summarize the basic idea by saying that we do not seek to judge quality of C2 products, but rather track variations in a decision maker's C2 products that we can attribute to the performance of an IT suite.

For a concrete example, if we want to understand how close a battlefield IT suite came to *perfection* at any (or all) times of particular interest in a battle, we simply compare the C2 products of the decision maker in question to those of the same decision maker operating in the same environment, always in the exact same situation as the subject, but now with a ground truth data stream. When the comparison indicates small differences, the IT performance is nearly perfect, and when the differences are large, the performance is far from perfect; if we want to know what was missed by the IT suite, we look into the details of the comparisons.

Similarly, we can compare a decision maker's battle command activity with a degraded network to that produced by the same decision maker using an optimal network. This comparative assessment is always grounded in the same situation, meaning the same mission details are understood at each time by the compared decision makers, except that the decision maker with optimal network resources functions with data provided via a network with a perfect message completion rate, i.e., with delays that reflect optimal network performance, through which the decision maker receives every message addressed to him. In this kind of comparison, a failure of the overall IT suite to gather data on some enemy entity due to the misplacement or absence of a sensor will not lead to any great differences in the comparison. This consequence occurs because the focus of the comparison is precisely on measuring differences in this decision maker's C2 products that are *attributable to* the network's performance and not on sensor management. Notice that in each example we described, we have deliberately avoided characterizing the quality of the cognitive activity of the decision maker with an improved data stream as brilliant, when in fact, it may or may not be; rather we have merely asserted that the judgment is maximal for the decision maker being assessed with respect to the improved data stream, though not necessarily maximum in general.

From the OISA method we described above, we believe that for a given military unit, its IT suite can be evaluated via objective assessors with absolute knowledge of the decision making agents they assess. Thus, by comparing assessments at various levels, we can identify places where a subordinate did the best he could, given the orders issued by his superior, and conversely, where the subordinate did poorly given those orders; in this method, the terms "best" and "poorly" are understood in terms of the data stream comparison that is being conducted. We note that the OISA method can only be implemented in the M&S context, since attempting this in a HITL experiment would require a never-ending query process of the C2 entity being assessed: what would you do now if we told you X-Y-Z was true? What now, if we told you A-B-C was true?

In the M&S context, implementing the described comparisons is not onerous. Having written the required software routines to model a decision maker at some level of the command hierarchy, we simply implement a clone of the agent and provide it an idealized data stream, and program the clone to act as an *objective assessor*, but not a participant in the surrounding simulation. The assessors compare their own decision inputs, processes and outputs to those of the assessed agent, and recording the comparisons in the form of metrics. The clone can provide objective assessments, in terms of any alternative data stream, because it possesses the same decision algorithm as the subject decision maker, a true reflection of absolute knowledge of its counterpart's reasoning processes. Note that in this simulation implementation, the objective assessor agent cannot command forces nor can it influence its counterpart in any way – its sole job is to assess the decision maker it shadows; in this aspect, the objective assessor agents are models of the OCs used at Army CTCs. Furthermore, the human analyst can choose which customized data stream to provide to the objective assessor agent, depending on the analysis foci; e.g., as described above, ground truth or perfect communications data could be provided, with which the objective assessor agent will form metrics expressing comparisons between its outputs and those of the simulation agents. In summary, the objective assessor agents are constructed to help the analyst objectively assess the performance of IT in supporting C2 within a simulation military unit. We conclude this section by noting that the OISA method has been implemented and used to assess notional IT systems (Davidson et al. 2008), and most recently employed in the dissertation work of Hudak (2008) and Hudak *et al.* (2008).

7. Conclusions and Future Work

In this paper we have identified M&S as an alternative tool for providing assessments of IT's support of C2, that overcomes various weaknesses present in HITL experiment methods. We observed that there are obstacles to modelling C2 entities within M&S environments, related to the massive C2 domain ontology and the issue of objectively measuring C2 products. We discussed how to overcome these obstacles through the identification of universal C2 products, such as SA and SU, that support all other C2 products, and the use of an objective C2 measurement method called the Objective Information System Assessment method. The next step in our work is to advance from concept development to execution, by applying the definition, representation and objective measurement choices described in this paper. We are developing further articles to describe details of the C2 entities modeled, the corresponding implementations of OISA using these C2 models, including the C2 metrics defined, and experimental designs, both analyses of alternatives and vulnerability studies, that are possible with such implementations.

8. Reference List

Alberts, D. S. and Hayes, R. E. (2003). *Power to the Edge: Command, Control in the Information Age*. CCRP Publication Series, Washington, DC.

Alberts, D. S. and Hayes, R. E. (2006). *Understanding Command and Control*. DoD Command and Control Research Program, Washington, DC.

Bernstein, R., Jr.; Flores, R.; Starks, M. W. (2006). *Objectives and Capabilities of the System of Systems Survivability Simulation (S4).* Final Report ARL-TN-260. U.S. Army Research Laboratory, White Sands Missile Range, NM.

Blechinger, P. I.; Kirin, S.; Riese, S.; Gorski, B.; Peters, D. (2004). *FCS Battle Command Analysis: Measuring the Contribution of Information to a Networked Enabled Force.* Technical Report TRAC-TR-04-016. U.S. Army TRADOC Analysis Center - Ft. Leavenworth (TRAC-FLVN), Ft. Leavenworth, KS.

Booz Allen & Hamilton. (1999). *Measuring the Effects of Network-Centric Warfare, Vol. 1.* Booz-Allen and Hamilton, Inc., McLean, VA.

Charles, P. and Turner, P. (2004). Capabilities Based Acquisition ... From Theory to Reality. *CHIPS*

Command and Control Research Program. Command and control bibliography. **2009**. URL: http://www.dodccrp.org/html4/research_c2.html (accessed Jan. 20,2009).

Crandall, B., Klein, G. A., and Hoffman, R. R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. MIT Press, Cambridge, Mass.

Darilek, R. E., Perry, W. L., Bracken, J., Gordon, J., and Nichiporuk, B. (2001). *Measures of Effectiveness for the Information-Age Army*. Rand, Santa Monica, CA.

Davidson, J., Pogel, A., and Smith, J. A. (2008). The Role of Battle Command in Information System Assessments. In the *Proceedings of the 13th Annual International Conference on Industrial Engineering Theory, Applications and Practice*. International Journal of Industrial Engineering, pp. 154-160.

Delbrück, H. and Renfroe, W. J., Jr. (1980). *History of the Art of War Within the Framework of Political History [Geschichte Der Kriegskunst Im Rahmen Der Politischen Geschichte]*. Greenwood Press, Westport, CT, vol. 4: The Modern Era.

Endsley, M. R. (1988). Situation Awareness Global Assessment Technique (SAGAT). In the *Proceedings of the National Aerospace and Electronics Conference (NAECON)*, vol. 3. IEEE, New York, pp. 789-795.

Gigerenzer, G., Todd, P. M., and the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press, New York.

Graham, H., Coppin, G., and Cummings, M. L. (2007). The PRIM: Extracting Expert Knowledge for Aiding in C2 Sense & Decision Making. In the *Proceedings of the 12th International Command and Control Research and Technology Symposium (ICCRTS 2007)*. U.S. Department of Defense, Command and Control Research Program, Alexandria, VA

Hewett, R. and Leuchner, J. (2003). Restructuring Decision Tables for Elucidation of Knowledge. *Data Knowl. Eng.*, *46*(3): 271-290

Hodges, J. S. (1994). Analytical Use of Data From Army Training Exercises - A Case-Study of Tactical Reconnaissance. *J. Am. Stat. Assoc.*, *89*(426): 444-451

Hudak, D. Assessing the impact of information on decision-making in contexts lacking well-defined utility functions. Ph.D. Dissertation, New Mexico State University, Dec 2008.

Hudak, D., Mullen, J., and Pogel, A. (2008). Determining the Impact of Information on Decision-Making in Contexts Lacking a Well-Defined Utility Functions. IJIE, Las Vegas, NV, pp. 102-108.

Jensen, D., Oates, T., and Cohen, P. R. (1997). Building Simple Models: A Case Study With Decision Trees. In *Advances in Intelligent Data Analysis. Reasoning About Data: Second International Symposium, IDA-97, London, UK, August 1997* (Eds. X. Liu, P. Cohen, and M. Berthold) Springer-Verlag, Berlin, pp. 211-222.

Klein, G., Phillips, J. K., Rall, E. I., and Peluso, D. A. (2007). A Data-Frame Theory of Sensemaking. In *Expertise Out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (Ed. R. R. Hoffman) Lawrence Erlbaum Associates, New York, pp. 113-158.

Kunde, D. Event prediction for modeling mental simulation in naturalistic decision making. Ph.D. Dissertation, Naval Postgraduate School, Monterey, CA, Dec 2005.

Kunde, D. and Darken, C. J. (2005). Event Prediction for Modeling Mental Simulation in Naturalistic Decision Making. In the *Proceedings of the 14th Conference on Behavior Representation in Modeling and Simulation*

Kunde, D. and Darken, C. J. (2006). A Mental Simulation-Based Decision-Making Architecture Applied to Ground Combat. In the *Proceedings of the 15th Conference on Behavior Representation in Modeling and Simulation*

Lamartin, G. F. (2004). Capability Based Planning - an Acquisition Perspective. In the *Proceedings of the MORS CBP Workshop*

Miller, N. L. and Shattuck, L. G. (2004). A Process Model of Situated Cognition in Military Command and Control. In the *Proceedings of the 2004 Command and Control Research and Technology Symposium (2004 CCRTS)*. U.S. Department of Defense, Command and Control Research Program, Alexandria, VA

Naylor, S. (2005). *Not a Good Day to Die: The Untold Story of Operation Anaconda*. Berkley Books, New York.

Network Systems Group. **2009**. URL: http://nsg.cs.princeton.edu/ (accessed Jan. 20,2009).

Newell, A. (1990). *Unified Theories of Cognition*. Harvard University Press, Cambridge, MA.

Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, N.J.

Russell, S. S., Dorsey, D. W., Ford, M. T., Campbell, G. E., Van Buskirk, W. L., and McCreary, A. F. (2006). Model Validation Is Not Simple, Even When the Model Is: Lessons Learned From a Computational Model of Performance. In the *Proceedings of the 15th Conference on Behavior Representation in Modeling and Simulation*

Starks, M. W.; Flores, R. (2004). *New Foundations for Survivability/Lethality/Vulnerability Analysis (SLVA).* Technical Note ARL-TN-216. U.S. Army Research Laboratory, White Sands Missile Range, NM.

Sun Tzu and Griffith, S. B. (1963). *The Art of War*. Oxford University Press, London.

U.S. Army. (2002). *The Objective Force in 2015.*

U.S. Army. (2003). *Mission Command: Command and Control of Army Forces.* Field Manual 6-0. Headquarters, Department of the Army,

U.S. Army. (2004). *United States Army Transformation Roadmap.* Office of the Deputy Chief of Staff, U.S. Army Operations, Army Transformation Office, Washington, DC.

U.S. Army. Army FY 2009 Budget Overview. **2009**. URL: http://www.asafm.army.mil/budget/fybm/FY09/overview.pdf (accessed Jan. 20,2009).

U.S. Department of Defense. (2002). *Mandatory Procedures for Major Defense Acquisition Programs (MDAPS) and Major Automated Information System (MAIS) Acquisition Programs.* Regulation 5000.2-R. Undersecretary of Defense (Acquisition, Technology and Logistics), Washington, DC,

Waters, M. L. (2004). From SOSO to High-Intensity Conflict: Training Challenges for FA Battalions. *Field Artillery Journal*: 28-35

Wickens, C. D., Gordon, S. E., and Liu, Y. (2004). *An Introduction to Human Factors Engineering*. 2nd ed.; Pearson Prentice Hall, Upper Saddle River, NJ.

Widder, W. (2002). Auftragstaktik and Innere Führung: Trademarks of German Leadership. *Military R.*, *LXXXII*(5): 3-9

Wiener, E. L. and Nagel, D. C. (1988). *Human Factors in Aviation*. Academic Press, San Diego.