SWEDISH DEFENCE
RESEARCH AGENCY

FOI-R—1034--SE
November 2003
ISSN 1650-1942

Scientific report

Per Wikberg, P-A Albinsson, Torbjörn Danielsson, Helena Holmström, Johan Stjernberger, Mirko Thorstensson, Malin Östensson

Methodological Aspects of Development and Evaluation of Military Command and Control Systems

SWEDISH DEFENCE RESEARCH AGENCY

Command and Control Systems
P.O. Box 1165
SE-581 11 Linköping

FOI-R--1034--SE

November 2003

ISSN 1650-1942

Scientific report

Per Wikberg, P-A Albinsson, Torbjörn Danielsson, Helena Holmström, Johan Stjernberger, Mirko Thorstensson, Malin Östensson

Methodological Aspects of Development and Evaluation of Military Command and Control Systems

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Command and Control Systems P.O. Box 1165 SE-581 11 Linköping	FOI-R--1034--SE	Scientific report
	Research area code 4. C4ISR	
	Month year November 2003	**Project no.** E-7080
	Customers code 5. Commissioned Research	
	Sub area code 49 Interdisciplinary Projects regarding C4ISR	
Author/s (editor/s) Per Wikberg P-A Albinsson Torbjörn Danielsson Helena Holmström Johan Stjernberger Mirko Thorstensson Malin Östensson	**Project manager** Per Wikberg	
	Approved by Erland Svensson	
	Sponsoring agency Swedish Defence Forces	
	Scientifically and technically responsible Per Wikberg	

Report title

Methodological Aspects of Development and Evaluation of Military Command and Control Systems

Abstract (not more than 200 words)

Due to the rapid development of technology and political changes, most defence forces are currently reorganising their tactics, tools and methods. Changes in thinking and practice in military command and control (C2) systems do not only include technical aspects but also terminology, culture, planning processes etc. These organisational changes are fast and thus methods and techniques for rapid evaluation are essential. Time is a crucial factor and the evaluation methods have to keep pace with the time cycle of the practical project. However, there is still the demand for reliable and valid results. Within the FOI-project Evolva, methodological principles and procedures for development and evaluation of military C2 systems are being developed and tested in cooperation with the Swedish Armed Forces. The work focuses on four essential parts of the test cycle: modelling as a tool for defining evaluation measures; environments for testing C2 concepts; tools, techniques and methods for data collection and tools, techniques and methods for instant feedback of test results. In this report, experiences from evaluation of military C2 systems are presented together with a number of cases to illustrate the results. The report also includes a presentation of some lines of future work.

Keywords

Further bibliographic information	**Language** English
ISSN 1650-1942	**Pages** 33 p.
	Price acc. to pricelist

Utgivare	Rapportnummer, ISRN	Klassificering
Totalförsvarets Forskningsinstitut - FOI Ledningssystem Box 1165 581 11 Linköping	FOI-R--1034--SE	Vetenskaplig rapport
	Forskningsområde 4. Spaning och ledning	
	Månad, år November 2003	**Projektnummer** E-7080
	Verksamhetsgren 5. Uppdragsfinansierad verksamhet	
	Delområde 49 Breda projekt spaning och ledning	
Författare/redaktör Per Wikberg P-A Albinsson Torbjörn Danielsson Helena Holmström Johan Stjernberger Mirko Thorstensson Malin Östensson	**Projektledare** Per Wikberg	
	Godkänd av Erland Svensson	
	Uppdragsgivare/kundbeteckning Försvarsmakten	
	Tekniskt och/eller vetenskapligt ansvarig Per Wikberg	

Rapportens titel (i översättning)
Några metodaspekter vid utveckling och utvärdering av militär ledning

Sammanfattning (högst 200 ord)
En snabb teknisk och politisk utveckling har lett till att många länder omorganiserar sina försvarsmakters mål, medel och metoder. De förändringar detta innebär för ledning, innefattar inte bara tekniska aspekter av ledningssystemet utan även terminologi, kultur, planeringsprocesser etc. Dessa förändringar sker i hög takt varför metoder och tekniker för snabb utvärdering är av högsta värde. Tid är en kritisk faktor och arbetet med att utvärdera de övergripande systemkonsekvenserna måste kunna anpassas till den takt med vilken förändringarna sker. Vid utvärderingarna gäller dock fortfarande samma krav på reliabilitet och validitet i slutsatser som i vanlig forskning. Inom FOI-projektet Evolva prövas och utvecklas metoder och tekniker för utveckling och utvärdering av militära ledningssystem. Arbetet sker i samarbete med Försvarsmakten och fokuserar på fyra grundläggande steg i en utvecklings- och utvärderingscykel: 1) Modellering för att definiera försöksdesign. 2) Datainsamling. 3) Försöksmiljöer. 4) Feed-back av försöksresultat. I föreliggande rapport redovisas erfarenheter från detta arbete tillsammans med ett antal exempel för att illustrera arbetet. Rapporten innehåller även en redogörelse för fortsatt inriktning av arbetet.

Nyckelord

Övriga bibliografiska uppgifter	**Språk** Engelska
ISSN 1650-1942	**Antal sidor:** 33 s.
Distribution enligt missiv	**Pris: Enligt prislista**

Contents

Executive summary

Introduction

Due to the rapid development in technology and political changes, most defence forces are currently reorganising their tactics, tools and methods. Changes in thinking and practice in military command and control (C2) systems do not only include technical aspects but also terminology, culture, planning processes etc. These organisational changes are happening fast and thus methods and techniques for rapid evaluation are essential. Time is a crucial factor and the evaluation methods have to keep pace with the time cycle of the practical project. However, there is still the demand for reliable and valid results. Within the FOI-project Evolva, methodological principles and procedures for development and evaluation of military C2 systems are being developed and tested in cooperation with the Swedish Armed Forces. The work focuses on four essential parts of the test cycle: modelling as a tool for defining evaluation measures; environments for testing C2 concepts; tools, techniques and methods for data collection and tools, techniques and methods for instant feedback of test results. In this report, experiences from evaluation of military C2 systems are presented together with a number of cases to illustrate the results. The report also includes a presentation of some lines of future work.

Modelling as a tool for defining evaluation measures

A relevant model, outlining relevant factors and sub processes, provides guidelines for what data to collect and how to interpret the results. The modelling activity helps to identify the critical parameters to be measured and also constitutes the language in common between practitioners and the system assessment staff. Modelling is thus a powerful tool to limit costs involved in data collection in large and complex systems. Modelling of complex systems aiming at evaluation should meet certain criteria: the modelling method must make efficient use of time and resources, the model must be empirically specified and it must be relevant to the context without being too complex.

Test environments

Military command and control systems can be studied in a variety of settings, both real and simulated. As an attempt to address some of the problems involved in conducting research on human aspects of command and control, two versions of a low-cost technique for conducting simulated command and control exercises using experimental designs have been developed and tested. The approaches, Experimental Simulation Exercise (ESE) and Computer Game Based Simulation, are presented.

Data collection

We focus mainly on a hypothesis driven approach, that is, to collect data about specified parameters identified from earlier analyses and modelling. However, we also acknowledge that all data collection is part of an intertwined blend of hypothesis testing and hypothesis generation. Irrespective of the focus of the specific data collection there are a number of standard methods that can be used for collecting process data. We describe a number of such data sources relevant to command and control. The phases of data collection: preparation, conduction and compilation are compared, also taking into account adaptivity and quality, two factors that can affect the level of technical support.

After Action Review

After action review (AAR) is a method used to provide and deliver instant feedback, originally developed for collective training exercises. This is done by collecting and presenting data that characterise the performance of the units in relation to preset goals and mission outcomes. Another area in which AAR concepts can be utilised is in the development process of new command and control systems. The basic approach is the same, where output from the involved participants provides an additional opportunity for data collection as well as for validation. The frame of reference for the AAR is the model specified in the process of modelling the problem. We present a model-based approach to present results from AAR to the involved practitioners, as an alternative to commonly used statistical graphs and charts. The basic idea is to use large paper sheets or video projectors to present results in a graphical model-based layout in order to force evaluators to interpret data according to the initially specified model. A framework called MIND is used to exemplify our AAR approach. The MIND framework has been developed at the Swedish Defence Research Agency (FOI) and has been used for AAR since 1995 (Morin, 2002).

Conclusions and future work

The aim of this paper has been to suggest a feasible way to conduct research in the process of developing C2 systems. The suggested focus in such a process is an iterative process of hypothesis testing in systematically designed tests. The approach has the potential to provide periodic or even continuous feedback rather than providing it only once. Some lines of future work are:

Extending the test environments

Based on terrain data, a virtual model of a part of a real exercise field is under construction. In this virtual model of actual terrain it will be possible to navigate freely, and to conduct double sided tactical exercises using commercial software. Our primary interest is to compare units conducting the same mission in a real environment and in a virtual environment. A series of tests in this matter will address training effects and perhaps even tactical applications at the soldier and squad level. Commercial software is also interesting from a C2 perspective. Combining simulated battle in commercial software with real C2 systems might both be a more resource effective and a more realistic setting for research on C2. The intended series of tests will thus make it possible to estimate strengths and shortcomings with such a research setting compared to settings based on units performing their missions in real environments.

Extending the data collection techniques

The project has shed light into future needs concerning the MIND framework. To date, the components in the framework handling communication focus on audio based radio communication. However, there are no extensive components that capture and present text based communication which is likely to increase in the on-going introduction of modern information technology into the command and control processes. Our intent is to develop capability to capture and present this type of communication in a tactical context. Work has been initiated to extend the framework in this aspect.

Extending the AAR Approach

The components in the MIND framework focus on providing efficient representations of raw data and useful techniques for handling them. There is, however, more than representation and handling of data in the analysis process. Collected raw data from a complex test setting are in a stable and accessible condition within the MIND framework, but they are largely separated from *the more important* overall insights and conclusions that evolve over time during analyses. These meta-data can be hard to find, recall, commonly share, and put into context again. They can lose their connection to their original settings. We are therefore developing a tool, called the *meta-data workbench*, for tackling this problem. The basic idea of the workbench is to provide a dynamic space for adding and coupling meta-data to original raw data. The workbench, together with all its meta-data, is stored together with all other MIND components and raw data. In this way, the meta-data are always connected to their original circumstances, and they are always accessible to the analysts involved. The objective representation of the situation can be combined and presented with a context based conclusion from a more in depth analysis rather than just presenting raw data.

We are also interested in finding and designing efficient representations in the MIND framework of abstract and dynamic processes, such as distributed decision processes, to be used for AAR as alternatives to the model-based spreadsheets. We are looking into the possibility to integrate the presentation of the results into the MIND-framework, an achievement that would fulfill and close the loop from modelling of complex processes to presentation of results and performance.

1. Introduction

Due to the rapid development in technology and political changes, most defence forces are currently reorganising their tactics, tools and methods. Changes in thinking and practice in military command and control (C2) systems do not only include technical aspects but also terminology, culture, planning processes, routines etc. These organisational changes are happening fast and thus methods and techniques for rapid evaluation are essential. Time is a crucial factor and the evaluation methods have to keep pace with the time cycle of the practical project. However, there is still the demand for reliable and valid results. Furthermore, the design task in the development of new C2 systems is normally too complex to be solved analytically. Still, complexity calls for a set of common theoretical models describing the system. Thus, the development process requires an evolutionary approach based on successive empirical tests. The test should be contrasted to the initial system models, evaluating and revising them and generating new hypotheses to be contrasted in additional tests. At the Swedish Defence Research Agency (FOI), methodological principles and procedures for development and evaluation of military C2 systems are being developed and tested in the project Evolva. The work, which is done in cooperation with the Swedish Armed Forces, focuses on four essential parts of the test cycle:

- Modelling as a tool for defining evaluation measures
- Different environments for testing command and control concepts
- Tools, techniques and methods for data collection
- Tools, techniques and methods for instant feedback of test results

This report presents and discusses experiences from research in the development and evaluation of C2 systems. The first section deals with modelling as a prerequisite for evaluation, as modelling is essential for understanding and describing the C2 process and its context. The second section describes test environments for military command and control systems. In the following section different methods for data collection, and related problems, are discussed. We then present methods for after action review, (AAR) and discuss problems regarding presentation of results from an evaluation. The final section provides a short conclusion and points at possibilities for future work.

2. Modelling as a tool to define evaluation measures

The initial modelling of the system to be studied is a crucial task in any empirical evaluation. Modelling is the construction of a model, based on data from the systems analysis. There are several definitions of the term ‘systems analysis’, but a definition usually involves some kind of procedure (more or less formal) for collecting and organising data about an empirical phenomenon. There is a variety of systems analysis techniques and approaches such as ‘task analysis’ (Annett et al., 1971; Drury et al., 1987), ‘job analysis’ (Harvey, 1991), ‘content analysis’ (Kolbe, 1991; Weber, 1990), ‘action analysis’ (Singleton, 1979), and ‘cognitive systems engineering’ (Hollnagel & Woods, 1983; Rasmussen et al., 1994). Despite the fact that these techniques differ somewhat when it comes to perspectives and procedures, they are rather similar. They are related to a scientific style of analytically approaching a certain phenomenon, in order to treat or analyse reality as a systematically connected set of elements (Gasparski, 1991). However, it is important to emphasise that systems analysis is not necessarily the same as evaluation or design. Systems analysis often refers to the initial study of complex phenomena in a defined setting; a kind of mapping of the most relevant factors from a specific point of view. A relevant model provides guidelines for what data to collect and how to interpret the results. Modelling is thus a powerful tool to limit some of the costs connected to data collection in large and complex systems.

Modelling of complex systems aiming at evaluation should meet certain criteria (Wikberg, 1997; 2001, Thorstensson & Wikberg, 2002; Wikberg & Thorstensson 2003). Firstly, efficient use of time and resources is an important aspect of modelling. One reason for this is that access to subject matter experts (SMEs), who are valuable participants in modelling sessions, is often limited. A complex system could involve several hundred possible issues for empirical evaluation but naturally there is rarely enough time or resources to conduct all these studies. It is important to realise that the specific research questions to be pursued will evolve during the developmental process (March, 1991; Strauss, 1988). Visions and ideas will change and be refined, models will develop from concepts to prototypes and the bases for subjective judgements will change over time (Waterson et al., 1997). In a development cycle, empirical research must keep up the pace with these changes. Consequently, a command and control evaluation study may have to be completed within merely a few weeks. Secondly, to enable evaluation, the model must be empirically specified, i.e. there has to be a definition of the set of empirical elements and the relations between them that corresponds to the model’s conceptual terms. Consider the example of a causal model (where A causes B) as shown in figure 1.

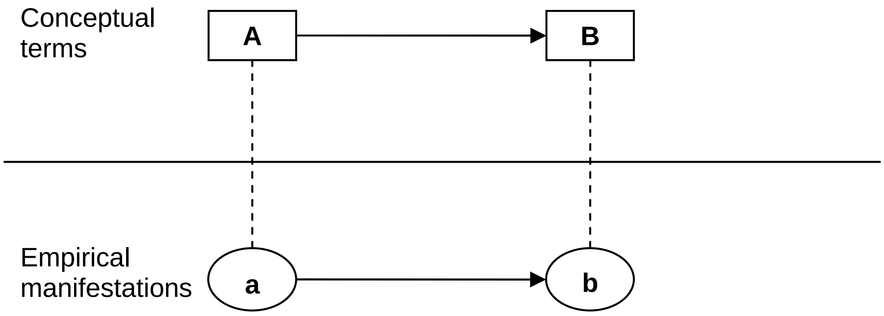

Figure 1. *Example of a causal model*

In order to measure and evaluate this model (figure 1), the following conditions have to be met (Wikberg, 1997):

- There must be a set of defined elements for representation:
  (1) A is defined as 'A'
  (2) B is defined as 'B'

- There must be a set of defined formal rules for relations between the elements which determine the possible combinations:
  A affects B (i.e. A is the most significant factor, of several other possible factors, affecting B)

- There must be a set of defined empirical elements and relations corresponding to the model:
  (1) ‘a’ is an empirical manifestation of 'A' (i.e. ‘a’ is the best, or most suitable, measure of 'A' within the context of the model)
  (2) ‘b’ is an empirical manifestation of 'B' (i.e. ‘b’ is the best, or most suitable, measure of 'B' within the context of the model)

Thirdly, the model must be relevant to the context without being too complex. Modelling of complex systems almost always produce complex models, as behaviour in the system varies between and within individuals and groups as well as over time. Process modelling of complex tasks may turn out to be a ‘never-ending story’, because of the need to model every specific chain of events, and data collection to support such an analysis may seem overwhelming. On the other hand, efforts to keep the model as simple as possible may result in insignificant or not useful models.

Case I: Modelling test design on effect of information structure on C2.

A study was performed in the context of the Swedish Armed Forces’ demonstrator program which aims at reshaping and adapting the defence from an invasion focused to a flexible network based defence (Swedish Armed Forces, 2002). The study, Demo 03, focused on how the C2 process is influenced by the information structure.

The research question of the study was analysed in a series of group modelling sessions with participants from the Armed Forces and FOI. In these sessions a hypothetical cause and effect relationship between information structure and C2 process was defined. The modelling comprised definition of relevant factors, how these are hypothetically related and how these factors were to be manifested in the exercise. These definitions form the basis for formulation of hypotheses and specification of how to measure the process. A graphical model illustrating the result of the modelling is shown in figure 2.

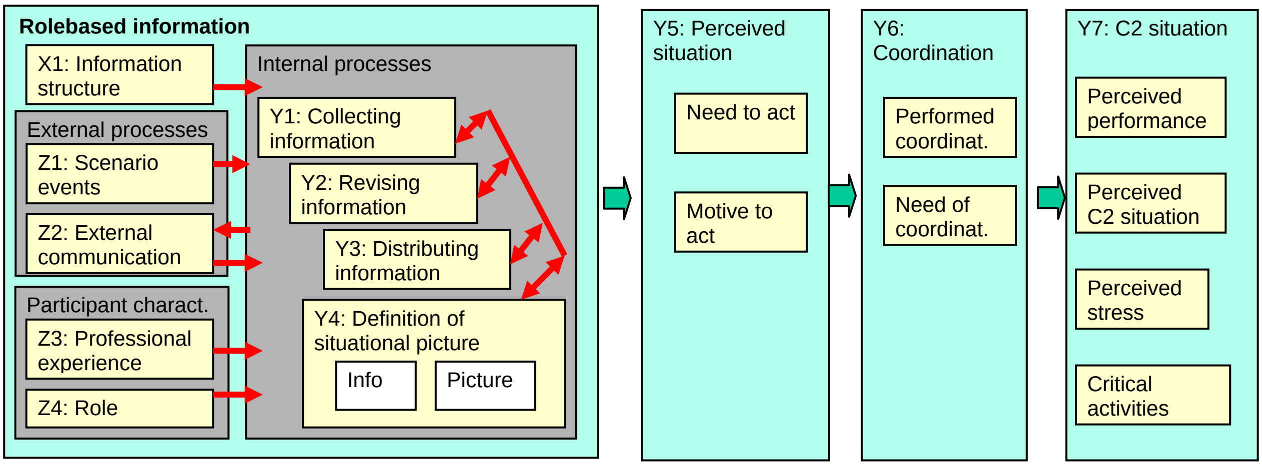

Figure 2. *The result from the modelling sessions where test design for Demo 03 was defined. The model describes a hypothetical cause and effect relationship between information structure and relevant factors of the C2 process. The model was used for formulation of hypotheses and specification of how to measure the process.*

The model describes how information structures **(X1)**, external processes and characteristics of the participants of the exercise **(Z1-Z4)** are related to a number of relevant factors in the C2 process **(Y1-Y7)**.

Another result of the modelling sessions was the actual exercise setup. The test was performed as a series of staff exercises conducted over a period of three days. The research manipulation consisted of a variation in ‘information structure’, X1, between these settings. Variation in ‘information structure’ was manifested through a variation in the access to information. The **Y**-factors were measured as the X1 setting was changed between each realisation. Each of the identified factors was empirically operationalised based on the definitions in the modelling sessions.

The effect of the variation between settings was tested with nineteen different hypotheses. An example of a hypothesis from the study is: *“If access to information is limited then the level of performed contacts between units to discuss coordination will decrease”.*

This particular hypothesis, which was rejected, was tested by comparing ‘level of contacts between units’ between settings. Data on the variable ‘level of contacts between units’ was collected by using:

1) Digital questionnaires where the participants estimated their need of coordination and described which contacts they actually had taken to discuss coordination.
2) Observers documenting which contacts the units had actually taken in order to discuss coordination.
3) Logging of voice and text communication between units. All communication was then classified on a number of factors. One of the classes was if the communication included discussions on coordination between units.

The hypothesis testing was facilitated by introducing a certain degree of automation in the testing and analysis process. This was made possible by creation of a database for handling the collected data in terms of recorded voice communication and text based messages intercepted in the C2 support system (IS-Mark). The content of the messages was filtered out, structured and presented to the assessment staff which could make the classification of the communication directly into the database. Afterwards the statistical analysis and hypothesis testing was facilitated by easy access to all the results divided into classes depending on the type of communication. The method and use of information technology as described, is now being refined and streamlined to be purposeful and efficient in assessment of this type of exercises.

Case II: Modelling of future decision processes in C2

In this study (Wikberg et al., in press) modelling was used to define alternative procedures for military decision making. The rationale for this study was that the realm of decision making in command and control today is wider than before and there is a general opinion that the traditional process of military planning is too slow. In addition, in developing the new Swedish command and control systems, the ambition is to incorporate new perspectives on decision processes such as “self-synchronisation” (Brehmer, 2002).

The study was conducted in two steps. The purpose of the initial phase was to define a basic theoretical model of decision making based on six theories of decision making from psychology and organisation theory. Together with subject matter experts, the theories were translated into ‘manuals’ like the “manual for appreciation” (“bedömandemallen” in Swedish) found in military regulations.

Theoretical perspectives on decision making in psychology

- Decision making as conditioned response on environment
- Decision making as rational choice between defined alternatives
- Decision making as dynamic process control
- Decision making as intuitive judgement based on expertise and recognition of contextual factors (Klein, 1993)

Theoretical perspectives on decision making in organisation theory

- Decision making as negotiation in order to achieve acceptable solutions
- Decision making as organic streams of events

The six different manuals were then merged into one model. The general features of this model are shown in figure 3.

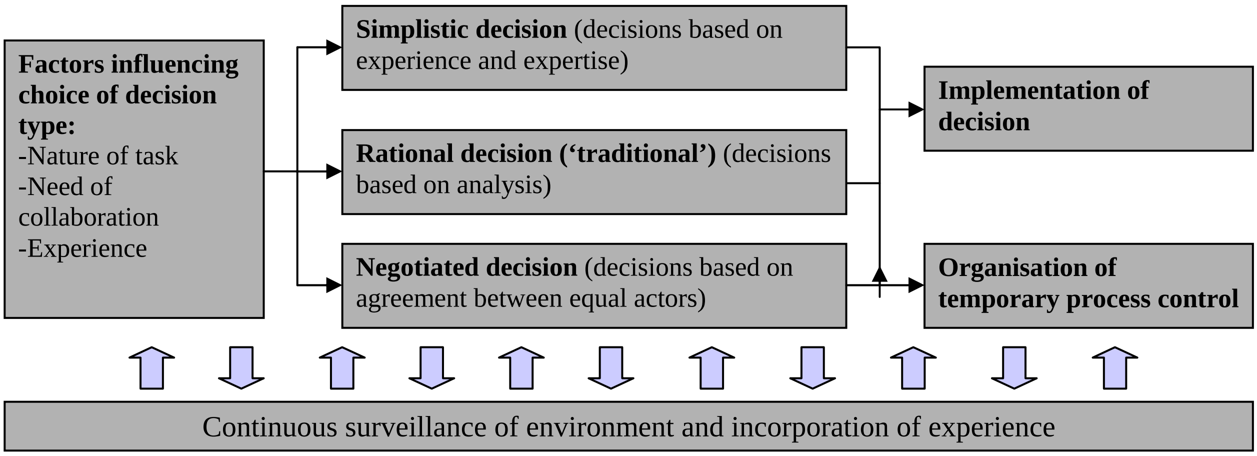

Figure 3. The *general features of the suggested decision making model described in Wikberg et al (in press)*

Note that figure 3 only shows the general features of the model. The model in detail is enclosed as appendix 1.

In the second step of the study, the model was used as a script in two war gaming exercises focusing on command and control. During the exercises participants described how they would act in the simulated situation. The process steps defined in the model were successively discussed based on the participants' different areas of expertise (se appendix 1 for a model in detail).

This example illustrates how an initial theoretical analysis (i.e. the modelling session) can be used to define ways to collect and organise data. In this particular case the model was used to manage the discussion during war gaming exercises but the same principle applies when defining what to measure in any empirical study.

3. Test environments

Human aspects of military command and control systems can be studied in a variety of settings, both real and simulated. Regarding decision making (which is a major research area within command and control) the classic research approach mainly involves laboratory settings. However, there has been widespread critique against this approach, for example in 1976 Neisser claimed:

> *"Contemporary studies of cognitive processes usually use stimulus material that is abstract, discontinuous and only marginally real. It is almost as if ecological invalidity were a deliberate feature of the experimental design." (Neisser, op cit., p. 34)*

On the other hand, using passive observation of occurring events in the field makes it almost impossible to draw any conclusions about the relation between stimulus and response, due to the lack of experimental control. Gist et al. (1998) discuss different contexts and research strategies concerning organisational simulation and define two dimensions which vary depending on the purpose of the simulation. The first dimension concerns the setting, i.e. the extent to which the game or simulation takes place within its 'real' environment. This dimension includes all research settings, from laboratories to actual field settings. However, a command and control exercise in the field is not necessarily more realistic in all aspects, than a command and control exercise in a laboratory (staff exercise facility). The second dimension deals with research strategy, i.e. the degree to which the researcher actively manipulates variables during a game or simulation. This dimension covers the spectrum of research strategies from the strict experiment with active manipulation of variables to observation studies. With regard to these two dimensions, eight types of games or organisational simulations can be identified (figure 4).

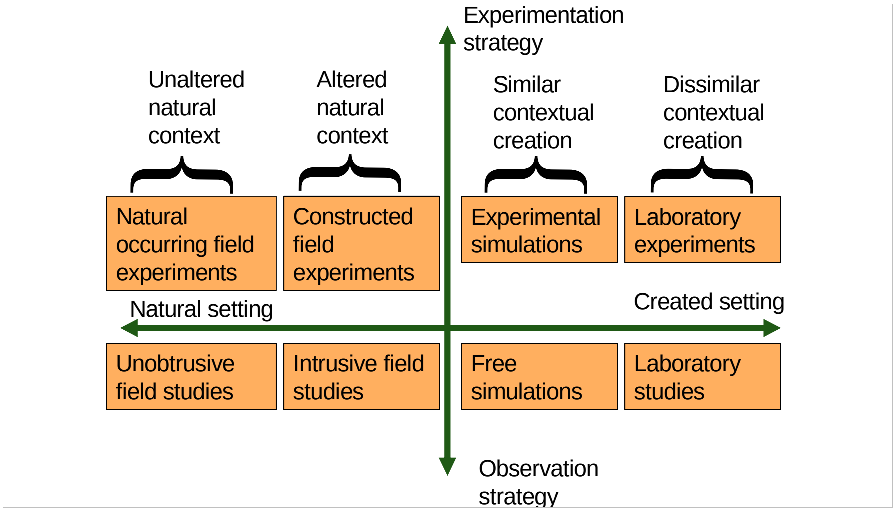

Figure 4. *Eight types of games or organisational simulations (Gist et al., 1998)*

Gist et al. (1998) stipulate that in debating the merits of laboratory versus field research the middle ground has been somewhat neglected in organisational literature. In some ways, experimental simulation can be said to represent that 'middle ground'. This methodology permits certain experimental rigour which is difficult to attain in the field, and offers contextual relevance, since measures are taken to simulate contextual elements of the natural, or real, setting. Nevertheless, designing a game for research purposes is always a delicate matter as the type of game directly affects the possibility to draw conclusions. There is also the constant struggle between purpose and research question and the resources available.

Choosing test environments when developing new command and control systems is an important issue. Resources for evaluation are often scarce and time schedules tight, but there is still the need for methods and techniques for fast and iterative evaluation. Evaluating human aspects of command and control processes in full-scale exercises is both difficult and time and resource consuming. For example, merely following and documenting one single event in a distributed decision making situation may require several observers and substantial technical equipment for registration. Another practical problem is limited access to exercise participants: interviews and other data collection activities may disturb the exercise. Moreover, evaluation of a complex system often calls for an iterative approach. When it comes to command and control processes this poses a problem, because 'natural' evaluation situations (i.e. full-scale exercises) occur infrequently and the objectives and focuses differ between exercises. As an attempt to address some of the problems involved in conducting research on human aspects of command and control, two versions of a low-cost technique for conducting simulated command and control exercises using experimental designs have been developed: Experimental Simulation Exercise (ESE) and Computer Game Based Simulation.

Experimental Simulation Exercise (ESE)

The concept of the ESE (Wikberg & Lundin, 2003) is best described as a simple, simulated staff exercise with an experimental design to enable the testing of hypotheses. Nonetheless, the ambition is to create a simulation that is realistic enough to produce data with direct practical implications. The scenario and input to forward the course of events that are used in the exercise are operationalised and designed to support the experimental design. The hypotheses typically concern abstract processes, for example how operators from different authorities value the same information in a command and control situation.

Basically, the ESE follows the same methodological approach presented in this paper with modelling, efficient data collection and an After Action Review (AAR). Users (military staff) play an important role in the ESE concept. They are involved in all phases of an ESE; problem analysis, the actual simulation exercise and the after action review that takes place right after the simulation. The problem analysis comprises a modelling session where the users take part in defining the terms and processes, generating hypotheses or research questions and a research design, operationalising the variables etc. Other preparations include creating (or choosing) a scenario and instruments for data collection (e.g. questionnaires or observation forms). The AAR requires efficient methods for data collection, since the ambition is to be able to present preliminary results as a basis for the discussion. The AAR also works as an additional opportunity for data collection, as well as validation of the results.

Case III: Distribution of information between C2 systems

An Experimental Simulation Exercise was conducted at the Swedish Armed Forces Joint Operational Staff's interaction lab (Wikberg et al. 2002). In this particular exercise, the aim was to examine processes involved in the administration of information between different technical command and control systems. The systems are still under development and at present it is not possible to transfer any information between them by technical means, why all processing and distribution of incoming information from sensors had to be done manually. During a two-hour modelling session with two military officers, the essential phases of a staff's information administration process were (hypothetically) defined. The research questions were:

- What limitations are there in processing and distributing different types of sensor information between current command and control systems?
- How do differences in designs and concepts between the current command and control systems affect situation pictures?

Using the software Visual Study 2002 (Dirsoft, 2000), a computerised questionnaire was designed based on the terms that were defined and operationalised during the modelling session.

Three command and control systems for the operational level ('IS Mark'/ground, 'IS FV'/air and 'LIM'/naval) and one tactical system for mechanised battalions ('SLB') were used in the simulation exercise. Additional technology support included staff support systems, MS Office, e-mail and telephone. Altogether, about sixty officers from all armed services, the Joint Military Intelligence and Security and five technical system operators participated. The design included three tactical staffs (one each for ground, air and naval) and one joint operational staff. The scenario described an attack of an enemy state which generated data from different sensors on the ground, at sea and in the air. Simulated sensor data were reported into each command and control system. Each staff then had to process and spread their system specific information to the other staffs and to simulated recipients on the tactical level in order to attain coherent operation and tactical situation maps. For each received input, the operator switched window on their desk top stations and answered the questions in the questionnaire.

Questionnaire data were compiled, processed and analysed successively during the exercise, in order to be able to present preliminary results in the AAR that followed immediately after the exercise. Preliminary findings were discussed with the participants and additional information and viewpoints were documented on-line. A view of the computer screen was projected onto the wall, so that everyone could follow what was being written down. A more thorough analysis of questionnaire and AAR data was conducted after the exercise.

Computer Game Based Simulation

Commercial game software is readily available and not very expensive. One possibility is to use commercial computer- and console games as low-cost simulators in experimental exercises. Even if computer games and consoles primarily are developed for entertainment, the low cost makes it an appealing alternative to ordinary exercises. It might be beneficial to use simulation technology when C2 requires extensive resources, have not been fully developed or is not available within the country. One area of application might be to use existing command and control systems for staff units to command units performing their tasks in a virtual environment. Using such an approach it might be possible, in an early stage and with a small investment in resources, to carry out research on how new technology will influence command and control methods. It is therefore of interest to develop a methodology for how commercial computer and console game can be used for different types of command and control studies and exercises.

Case IV: Commercial Software in a Ranger C2 Exercise

A Computer Game Based Simulation was conducted at Norrlands Dragoon Regiment. The aim was to explore the possibility of using commercial game software in an exercise where command and control are evaluated. Commanders' use of real-time information from Unmanned Aerial Vehicles (UAV) was studied using commercial software to simulate the tactical context of command and control. The question is relevant in the development of a network based defence as this is expected to generate large amount of real-time information from the battle space. Will commanders with access to detailed real time information interfere in operations on lower level, thereby loosing their general perspective?

At the regiment, ranger battalions are trained for combat and reconnaissance missions in a large variety of environments. Information transfer within a ranger battalion is limited compared to other units since a ranger task force commander can not bring extensive equipment and analysis tools, i.e. the ranger must be able to carry his combat load. Consequently, vast amounts of information from sensors such as UAVs, satellites and integrated helmet and display sight must be analysed by a battalion staff deployed at distance from the target area. Consequently, if the ranger battalion is equipped with, or supported by, such sensor systems, the rear command post will have better access to real-time information about the situation in the target zone than the task force commander. A number of hypotheses concerning effects of the staff's access to real time information on command and control methods were tested.

The study was arranged as an Experimental Simulation Exercise with one battalion staff and three ranger units. The ranger battalions' ordinary tactic command and control systems were used to command the different task forces. The tasks were carried out in a virtual commercial PC game environment. From one of the ranger units the battalion commander had access to real-time information from an UAV and from helmet assembled cameras. There were no such information from the other two ranger units. The structure of the exercise is illustrated in figure 5.

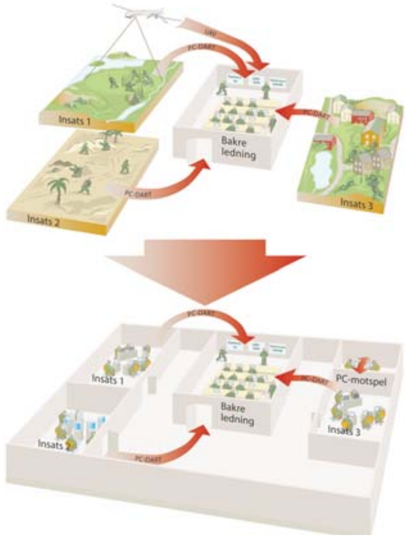

Figure 5. *The command and control conditions represented in virtual environments*

The participants were four career officers and 24 regular officers from the ranger battalion. There were also two technical officers from the Swedish Army Technical School (ATS). All participants were men. The participants were divided into one ranger headquarters and three task forces. Each of the three task forces consisted of one career officers and six to nine regulars from ranger battalion 04. The staff was composed of three career officers and two soldiers.

Three local networks (LAN) were organised, see Figure 5. The staff used their regular tactical system "IS MARK" while the task forces performed their missions in three virtual environments. A commercial, non-modified, game was used in each LAN. The PC game enabled the soldiers to integrate their actions through their personal terminals. In two of the LANs the computer generated opponent provided by the software was used as enemy. In the third LAN the enemy was controlled by other soldiers. Communication between the staff and the three task forces, i.e. LAN 1-3, were made through the ranger battalion's ordinary communication device PC DART. The tactical situation was followed up by the staff and registered in IS MARK on maps generated from the game. Each LAN was structured as following:

LAN 1 (Task force 1) was made out of eight computers and the game Delta Force: Task Force Dagger from Novalogic. The force was assisted by an UAV which circulated the area of interest maneuvered by the staff. The staff had a screen showing the image from the UAV. The computer screen of the commander of the task force was filmed and the image from the camera was directly available to the staff. The artificial intelligence served as the enemies

LAN 2 (Task force 2) used six Xboxes from Microsoft and the game "Ghost Recon" from Ubisoft. The artificial intelligence served as the enemies.

LAN 3 (Task force 3) consisted of eight computers where the game "Roque Spear" from "Red Storm Entertainment" was used. The enemy was two additional soldiers.

The staff was in charge of the units and worked in accordance with ordinary procedures. The participants practiced in a tutorial generated by the game in order to get necessary basic gaming skills. Communication and tactic were also trained as well as handling PC DART. Different scenarios were used for the pre training and the ordinary experimental training. Task force 2 and 3 executed the operations three times and stand-by force 1 did their operation two times. After execution each stand-by force did their debriefing as normal procedures.

Data was collected using observers, questionnaires and technical registration of communication. Raw data and preliminary results received from the data analysis was presented and discussed with all participants of the exercise in an AAR.

4. Data collection

Empirical evaluation always involves collecting empirical data. As we see it, the data collection is both hypothesis driven and hypothesis generative. Thus, the data collection will always be a mix of a modelling procedure and an empirical evaluation task. In this paper, we focus mainly on a hypothesis driven approach, that is, to collect data about specified parameters identified from earlier analyses and modelling. To collect data without an idea of what we want to measure would be a less cost-effective and more time-consuming strategy. But as always, there is a trade-off; the advantage of not having to filter large data sets has to be balanced against the disadvantage of not being able to use the data to study other phenomena than the ones specified in advance. Furthermore, even though we might be especially interested in a particular set of data, hypothesised to be a point of measure for a particular manifestation of certain phenomena, these data might be hard to interpret without a significant amount of contextual information. In turn, the need for such contextual information calls for more sources for the data collection. Therefore, especially for complex tasks such as command and control, we often have to adopt an exploratory strategy and thus collect data about many parameters that could be of interest. Woods (1993) suggests *process-tracing methods* to tackle these inherent challenges. Using process-tracing methods "does not mean that rigor or control […] must be sacrificed. Rather it means using a wide set of research tools to converge on an understanding of the phenomenon in question." (op cit, p 251). Thorstensson and colleagues (2001a) give examples of data sources that are useful for the command and control domain (table 1):

Table 1. *Data collection registration methods*

Data collection method	**Significance**
Observation	Observers' manual recording of organisational activities. All observer reports specify time (hour and minutes) and place.
Recording of communication	Automatic digital recording of radio and telephone communication. All communication is labelled in terms of time.
Registration of positions	Automatic registration of geographic positions using satellite navigation equipment
Photography	Documentation of activities and situations using digital camera. All photographs must be manually recorded and labelled in terms of time.
Video recording	Video recording of activities and situations.
Event cards	When there is an evaluation situation and some predefined events are expected to happen to certain individuals, these are equipped with event cards. For example, extra personnel serving as casualty markers might be equipped with causality cards on which they make notes during the exercise (Thorstensson et al., 1999). They note the time for specific events in the medical attendance.
Collecting log files	When support systems are used, log files that contain interesting parameters are collected from these systems. This facilitates the study of system usage and availability, in the overall context.
Collection of (evolving) tactical documents	Documents belonging to orders, reports and briefings etc are a rich source of information. Copies of these (over time as they evolve) will contribute to the exploration of the information situation in the command posts and the field units.
Registration of table for recourses and activities	The rear echelon command posts often use tables for presenting situations and prognoses regarding important resources and key activities. Documenting these tables by regular documentation, photographs or other means available enables the analysis of the information situation and activities in the command post.
Questionnaires	Personnel's thoughts, opinions and feelings can be registered using questionnaires.
Interviews	Interviews can be used to get a picture of how personnel perceive certain matters.
Physiological measures	Some physiological measures, such as hormone levels or blood pressure, can be used to indicate for example fatigue and stress.

Irrespective of the main focus of the specific data collection, hypothesis testing or hypothesis generation, there are a number of standard methods that can be used for collecting process data. All the methods for data collection listed above are adaptable to different levels of technical support (Thorstensson et al., 2001a). We have used different sets of methods in previous field trials (Morin, 2001; Crissey et al., 2001; Thorstensson et al., 2001b; Wikberg, 2001) and also in the simulation exercises described in this paper.

The phases of data collection can be separated in to preparation, conduction and compilation. Two factors, which also affect the level of technical support, are adaptivity and quality. Using these phases and factors, we can compare manual with automatic data collection. The low scores for automatic collection are preparation and adaptivity. To automate the work involved in the preparation phase or to automate the ability to adapt to changes in the world would require a complex automation system to handle the ill-structured and irrational nature of this type of work. Manual data collection would handle these situations better. For conduction, compilation, and quality, however, a successfully implemented automatic approach would outdo humans in efficiency. Though, a prerequisite is that the data are directly accessible and interpretable by the technical system.

In a one-shot limited field trial, manual methods are often the most appropriate. In a series of evaluations, or in large-scale evaluations where the results from data collections need to be presented shortly after the collection, automated methods are often preferable. However, in many cases there are no alternatives to the resource-demanding manual methods. Only small portions of a distributed decision can be registered with technical systems. Even with access to video registration of a staff process, the video data normally have to be interpreted manually.

Implementing manual methods using observers is rather straightforward, but skilled SME observers are rare and therefore it may be necessary to use novice observers. Using novice observers requires thorough instructions and training to achieve qualitative results (Jenvald et al., 2002).

Even with access to SME observers, such as staff instructors, there are difficulties in collecting and interpreting the observations. As Bainbridge et al. (1993) note: "*expert descriptions of complex and dynamic processes are often used but seldom analysed*". There is a risk that rational or mythological 'operator logic' may evolve that does not correspond to reality (Rasmussen, 1993). Furthermore, documenting experiences and data is a full-time task. Often a staff instructor has other obligations, such as coordinating the exercise, which limits the possibility to get data that are useful for evaluation. Data collection and evaluation should be the primary tasks of the observers (Wikberg, 2001).

Case V: Data Collection in an Emergency Response Exercise

To exemplify the data collection process, we will briefly discuss a study of an emergency response exercise staged in central Florida, conducted as a joint effort by Swedish and American researchers (Crissey et al., 2001). They used various techniques to collect data from the unfolding events. The site of the exercise was located just north of the University of Central Florida (UCF) Campus. Orange County Fire and Rescue and Seminole County Fire and Rescue jointly handled the incident with assistance from Orange County and Seminole County Sheriff departments and UCF Campus Police.

The exercise was based on a scenario where a container on a pickup truck starts leaking a suspicious substance in a parking lot outside a shopping mall. The substance, possibly chloride, must be identified before the exercise can unfold. Casualties, if any, must be found, removed from danger, and treated. The involved agencies must each make decisions specific to their roles and perform accordingly. Because of the complexity of the exercise, the assumption of responsibility and the leadership role was to be passed from agency to agency as the exercise progressed. The responding task force included two hazardous materials units and a mobile command post for incident command. As the scenario unfolded, the leak was positively identified as chlorine, and victims were reported. Special teams and medical aid units got ready to extricate the victims and to contain the chlorine leak.

Figure 6. *Fire fighters extricate a victim from the contaminated area (Photographer: Johan Jenvald)*

During the exercise the research teams and subject matter experts collected data using several different sources. Table 2 shows an inventory of the information sources that were used in the Orlando exercise and what topics the various sources covered. One observer monitored the tactical radio network and entered text comments as the exercise progressed. The other observers followed the exercise in the field. They used digital cameras together with structured reports to document the unfolding events. Throughout the exercise the GPS receivers tracked the position of the rescue vehicles, and radio traffic was recorded. When the exercise ended, the observers gathered in a data collection central where they compiled the various data into a common framework (the MIND framework). Position data from GPS receivers were downloaded and converted to the appropriate coordinate system, the digital photographs were downloaded and annotated using the structured reports that accompanied them, and observation reports were digitized.

Table 2. *Inventory of information sources used in the Orlando exercise (Crissey et al., 2001)*

	Information sources				
Topics	**GPS**	**Radio Networks**	**Casualty Report Cards**	**Photo-graphs**	**Manual observation**
Time	●	●	●	–	●
Unit positions	●	–	–	–	●
Unit deployment	–	●	–	–	●
Unit activity	–	●	–	●	●
Unit information status	–	●	–	–	●
Casualty type	–	–	●	●	●
Casualty treatment	–	–	●	●	●
Casualty location	–	–	●	–	●
Casualty status	–	–	●	–	●
Chemical substance	–	●	–	●	●
Terrain	–	–	–	●	●
Weather	–	–	–	●	●

Legend: ● information source is used; – information source is not used.

5. Feed back of results using After Action Review

After Action Review (AAR) (Rankin et al., 1995) is a method introduced some 30 years ago that has been developed ever since to provide and deliver feedback after a collective training exercise in order to enhance training value. This is made possible by collection and presentation of data that characterise the performance of the units in relation to preset goals and mission outcomes. The process of AAR seeks answers to the founding question "how did the unit do?" The answer may not be immediately obvious to the participants or to those who planned and controlled the exercise. Morrison & Meliza (1999) explain how this understanding is easily lost or obscured by "the fog of war" or just the simple fact that it is impossible for each and everyone to overlook the amount of distributed, but related, events and action going on during a collective training exercise.

An important role of the technological post exercise support is to objectively present context based action as a way of "fast forwarding" into the questions and phase of reflection. An important part of the AAR is thus the multimedia representation of distributed tactical events. This is made possible by adequate data collection and analysis of predetermined performance parameters and other data that represent the mission, such as unit movements, communication events, geographical data, weapons effect, orders given, and so on. Figure 7 gives an example of how the different data can be processed and presented in a technological support system for AAR. Basically, the support system shows the flow of events and units (as movements and positions on geographical maps), snapshots of certain interesting events or actions (as annotated photographs, videos and textual observations), and communication (as speech from radio networks or as text from computerised communication). All data are synchronised and connected to a timeline used as a basis for reconstruction of the exercise.

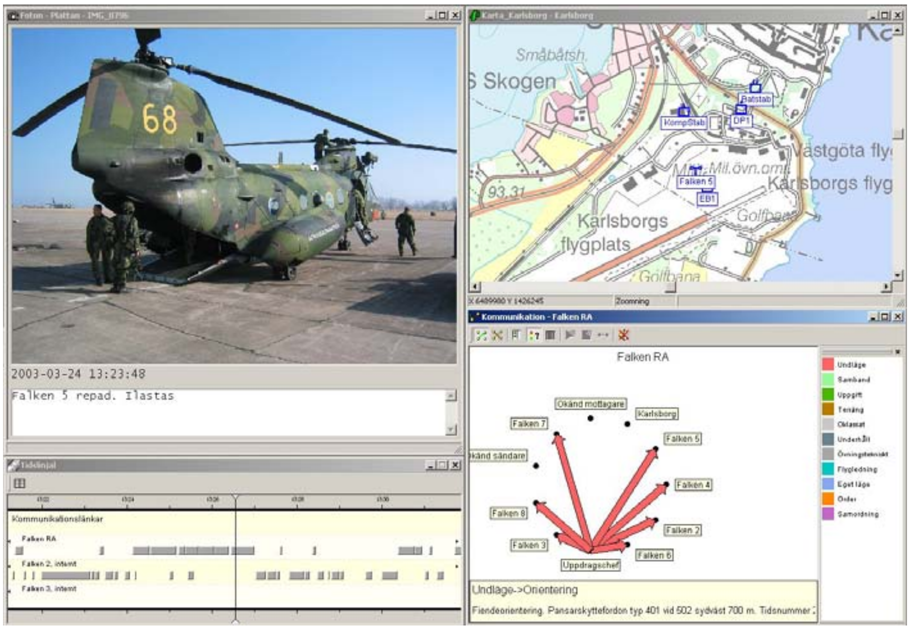

Figure 7. *A multimedia representation of a distributed tactical helicopter operation presented in the MIND framework.*

The traditional use of AAR methods and technology, described above, is to structure training and evaluate exercises in order to achieve or increase training value. Another area in which AAR concepts can be utilised is in the development process of new command and control systems as the AAR focus on rapid and iterative evaluation. Our experience is that the time frame from problem analysis to presenting results is between one week and six months depending on the problem (Wikberg & Thorstensson, 2003). Thus, the normal procedure of research reports is often too slow to keep pace with the development process. It must be possible to "deliver" the analysis of a test in time for the decisions on the progress of the development project. An AAR is then an opportunity to present raw data and preliminary results in a comprehensive way in order to discuss them with the participants as soon as possible after a test.

One important issue in the development of new C2 systems is that highly abstract processes such as decision making, situation awareness etc. must be analysed and developed. Consequently, data on these matters must be collected, analysed and presented in an AAR. This data collection and analysis comes with inherent problems.

In force-on-force it is easy to tell whether a mission has been successful or not, but it is more complicated to identify and measure the critical parameters of failure or success of the decision process or other non-technological contributions, such as leadership and how operators perceive the performance of the system. In most cases objective measures of performance are available but the subjective ratings of effectiveness have to be obtained from staff instructors, superiors, peers or subordinates judging the level of the staffs efficiency, the participants perception of a situation, perception of work load or stress, perceived difficulty, risk perception etc.

Another problem when abstract command and control processes are discussed in an AAR is that the time available for data analysis is limited. Consequently, data from decision processes are normally limited to quantitative data for statistical analyses. Presenting the results from such analyses poses a major problem as the participants of the exercise often has a limited knowledge and understanding of statistics. The evaluation often becomes a lecture on statistics rather than evaluation of a decision process (Wikberg & Lundin, 2002a). The time available for presenting and discussing the result is normally about one hour. Consequently, it is important to find procedures to present statistics in a straightforward way.

We evaluated the "normal" technique to present statistics in the format of graphs in the context of an 'Experimental Simulation Exercise' based on the results concerning the involved experienced officers (Wikberg & Lundin, 2002a). Two factors that effect the perception and interpretation of the statistical representations were tested. The first concerns cognitive dissonance, i.e. the difference between what result the individual expected and what was actually attained. If there is a great difference between these, assimilation may be difficult. Secondly, the design of the graphs as well as the verbal presentation of them may be indistinct or even incorrect. Two additional inquiries were distributed to the participants. The first, distributed directly after they had answered the regular inquiry concerning the problem studied in the ESE, was designed to survey to what degree the participants understood what comparison the graphs intended to represent. The second, distributed directly after the presentation of the results from the regular inquiry, was designed to elucidate which results the participants believed they would attain. The results indicate that the foundation for the AAR was incorrect since the understanding of the results only was 41 %. This could partly be explained by the fact that there was a difference between expected and attained result. Consequently, the statistical results used in an AAR must be presented in a context based and understandable manner to the participants.

We have tried out ideas on how to tackle this problem based on presenting a model-based evaluation of the abstract phenomena in the context of the complex system. The basic idea is to use large paper sheets or video projectors to present results in a graphical model-based layout in order to force evaluators to interpret data according to the initially specified model (Wikberg & Lundin, 2002b). This spreadsheet can also be used during the data collection part of the evaluation, forcing the observers to sort their data into the studied model. The model and its definitions form the basis for the definition of data collection. Documentation and interpretation can be conducted successively during frequent evaluation meetings. This approach makes it possible to present a documented result when the test is finished as opposed to spending days or weeks to analyse the data before being able to present any results.

Case VI: AAR in a model-based evaluation of information support

A study was performed focusing on the cooperation between different levels in the chain of command in an army division (Wikberg & Lundin, 2002b). The purpose was to study roles responsible for supplying the system with information in accordance with the succecsively identified need of information. The problem in focus of the study was structured in a modelling session including personell from the Army and the Swedish Defence Research Agency. The purpose of the modelling was to identify the most relevat factors of the problem in focus and how these factors should be measured in the given context of the Armys staff and communication exercise.

The resulting model was used as an observation protocol by military staff instructors. Following the staffs the staff instructors documented identified events, incidents procedures etc. that corresponded to the defined measures on site. Every day an evaluation meeting was held where each staff instructor presented his documentation.

Figure 8. *A model based After Action Review (AAR)*

All data where related to the factors of the model by using a print out of the model on a paper sheet of approximately 100 vs. 250 centimetres (See figure 8).
Each documentation was included in the model as raw data. A new version of the model was printed for the next meeting. After a couple of days of data collection, the documented data were used to draw some preliminary conclusions on a heuristic basis on each of the factors identified in the modelling session. These conclusions on factors where summoned into a more general discussion. The resulting graphical model based documentation is shown in figure 9. The spreadsheet contains the same information as an ordinary test report (including an introduction, sections on method and results and conclusions).

Consequently, the study contained an AAR each day where raw data were presented and analysed. After a couple of days, preliminary results were also presented and discussed. The major advantage of the approach was that the primary recipients of the results were engaged in the AAR sessions. This meant that they had first hand access and knowledge of the results and the context of the study in advance of the documented report.

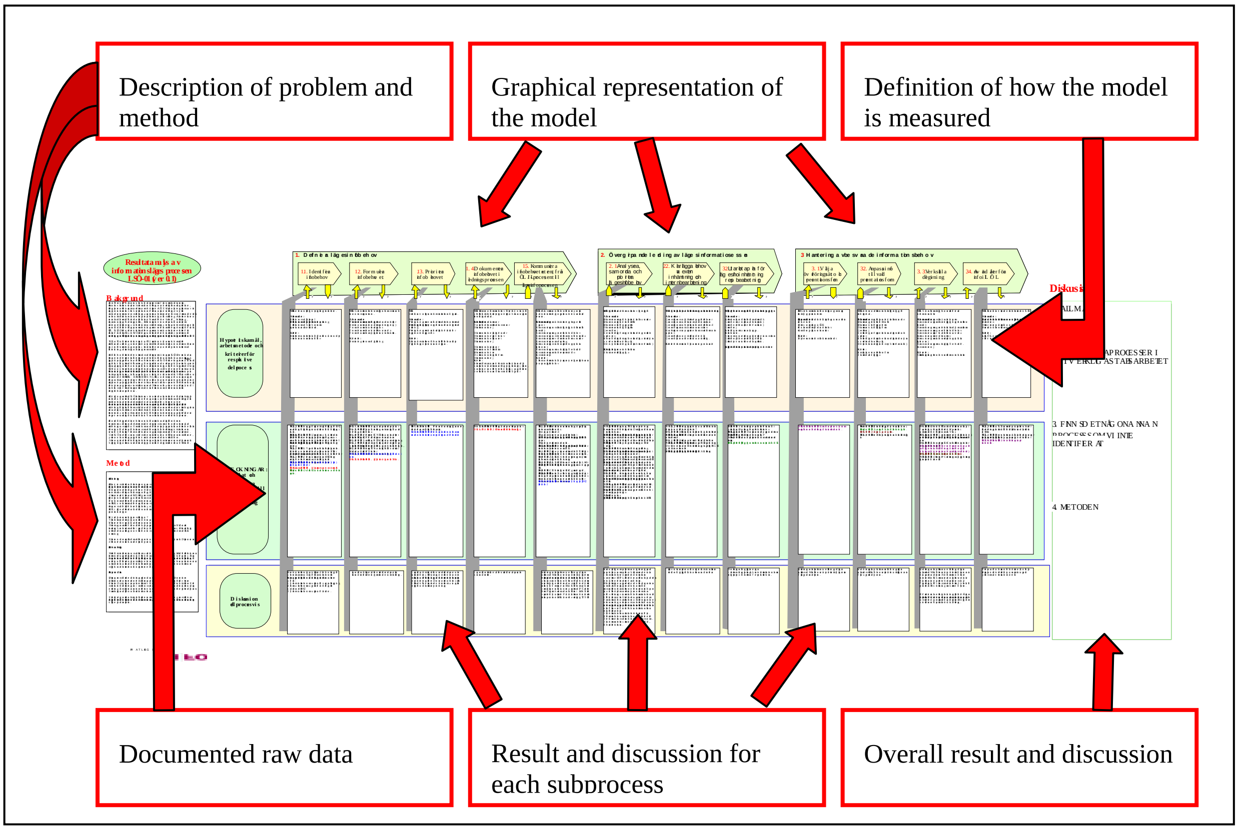

Figure 9. *The graphical layout of the documentation of a model based evaluation. In this particular case we used a paper sheet of approximately 100 vs. 250 centimetres why it is not possible to read the text in the figure shown here.*

6. Conclusions and future work

The aim of this paper has been to suggest a practicable way to conduct research in the process of developing C2 systems. The suggested focus is an iterative process of hypothesis testing in systematically designed tests for the purpose of system development and continuous task and organisational design. In the case of new C2 systems this is a crucial feature as there is reason to believe that it would be premature to specify a program in detail in advance.

The iterative design process of formulating hypotheses, contrasting them in various environments, turning back to the initial model for evaluation and revision, and generating new hypotheses to be contrasted in additional simulations has the potential to provide periodic or even continuous feedback rather than providing it only once. Consequently, it should be possible to define sub goals as the design activities progress instead of focusing the development on a program idea with an early defined and fixed set of demands on the system.

Extending the data collection techniques

So far, the components handling communication in the MIND framework have been focused on collecting, presenting and analysing analogue radio communication and other audio based types of communication. However, we have limited capability to capture and present text based communication. In the future, text based communication such as e-mail and other forms of digital transmissions are most likely to become the predominant way of communicating. Our intent is to enhance the capability to capture and present this type of communication in a tactical context and also to integrate the component as a part of the MIND framework.

Extending the test environments

Our work with test environments based on simulated environments in commercial software will continue. Focusing on commercial software already available at a low cost in every software store is basically a matter of resources. The difference in level of investment between a commercial product and tailor made software is huge, to say the least. There is also for every type of unit, a need for a set of simulated environments, calling for a number of different software, to make it possible to choose product depending on the purpose and nature of the test or exercise. Furthermore, there is no need for system maintenance using commercial software as new products continuously enter the market.

The experiences from the earlier mentioned test at the Norrlands Dragoon Regiment, in which the battalion staff used the ordinary command and control system to control three separate ranger units performing their missions in the virtual environments of commercial software in three local networks, will form the basis for further tests. Based on terrain data a virtual model of a part of the regiment's exercise field is under construction. In this virtual model of actual terrain it will be possible to navigate freely and to conduct simple double sided tactical exercises using the commercial software Battlefield 1942 (EA Games, 2003). Our primary interest is to compare ranger units conducting the same mission in a real environment and in a virtual environment.

A series of tests in this matter will of course address training effects at the soldier and squad level. Are the simulated environments of commercial software realistic enough to train squads in tactics in the same principal way simulators have been used in the Air Force for decades? Tactical application is also perhaps possible. It might be appropriate for a unit to rehearse a mission in a virtual environment which is a model of the actual terrain of the mission.

But, commercial software is also interesting from a C2 perspective even if the commanders at higher levels do not navigate themselves in the virtual environment. This principal situation is obviously similar to the one in a real battle where it is not possible for a commander to be on every location on the battlefield. Thus, there is a need for effective C2-systems where commanders can manage the distributed units at distance. Thus, combining simulated battle in commercial software with real C2 systems might both be a more resource effective and a more realistic setting for research on C2. The intended series of tests will make it possible to estimate strengths and shortcomings with such a research setting compared to settings based on units performing their missions in real environments.

Extending the AAR Approach

The components in the MIND framework have, to date, focused on the task of making various types of raw data graspable, coordinated and ready-to-work-on. The framework provides data presentation views that aim at making the data as informative as possible, considering the goals of the analyses. Furthermore, it provides tools for handling the data, such as navigating, exploring, filtering, and synchronizing. There is, however, more to data analysis than efficient *representation* and *handling* of data. When conducting analyses using MIND, the products of such analysis work include insights, reflections, questions raised, hypotheses and the like. Such products of course give further input into the analyses, guiding the analyst among questions and possible answers. The products are data themselves, on a meta-level. These meta-data have always been stored outside the MIND framework, as scribbles in a notebook or figures in a spreadsheet, and often stayed more or less as personal reflections for each analyst working with the data. The raw data are in a stable and accessible condition, but they are largely separated from *the more important* overall insights and conclusions that evolve over time. This has several drawbacks. Especially if there are several analysts working together with the data. The meta-data can be hard to find, recall, commonly share, and put into context again. They can lose their connection to their original settings. Ultimately, after reporting the results from the analysed data and delivering the conclusions to the customer, the meta-data connections to raw data might fade away behind concluding statements and generalisations.

We have acknowledged this problem and have developed a tool, called the *meta-data workbench*, for tackling it (Figure 10). The basic idea of the workbench is to provide a dynamic space for adding and coupling meta-data products to original raw data. The workbench consists of a timeline that covers the whole time period of the collected data. The analyst can, by clicking on appropriate points of time in the timeline, add meta-data *notes* of situations of interest. The notes comprise a timestamp, a caption, a description and an icon. The icons and captions are shown in the timeline, and the descriptions are reached by clicking the icons. Furthermore, each note is coupled with the current arrangement of the underlying views, tools, and associated data. This means that we can adjust, arrange, open or close the MIND components to reflect as clearly as possible the identified situation of interest, and then connect this configuration to the note added. In this way, when later investigating a note, we can activate the proper configuration by clicking the note. The notes can be arranged in *note fields* to, for example, group them according to type, importance, level of abstraction or any other criteria of classification. The note fields can be minimized or maximized as appropriate.

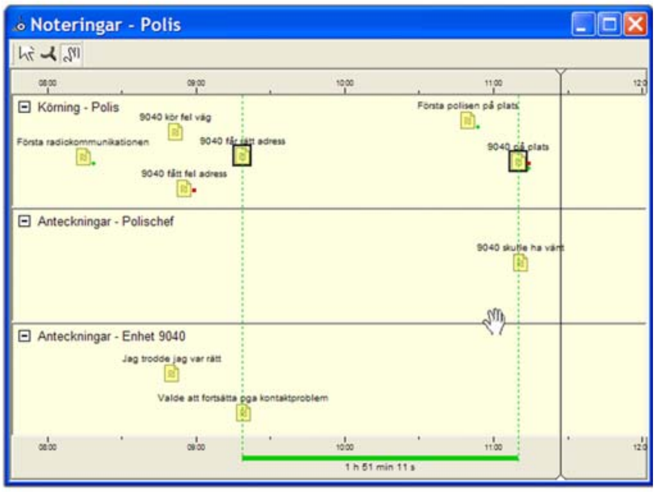

Figure 10. *Screen shot of the MIND Workbench Tool. This example workbench comprises three note fields: police in general, police commander, and police unit 9040, where the first field contains notes of the analysts, and the other contain notes from involved actors. The police commander has commented on the notes of the analyst, and the police unit discussed in the notes, has himself commented on these notes. The current replay time is shown by the horizontal line cursor. In this example we have selected two notes to get the time span in between.*

The workbench, together with all its meta-data, is stored together with all other MIND components and raw data. In this way, the notes are always connected to their original circumstances, and they are always accessible to the analysts involved. The notes thereby function as landmarks that highlight and draw data together. This encourages analysts to consider the meta-data as naturally as the raw data, thus being able to swiftly switch between investigating the meta-data conclusions or hypotheses and reaching new conclusions or generating new hypotheses from the raw data. Furthermore, the dynamic nature of the workbench tool makes it possible for involved practitioners or subject matter experts to comment on the data by just adding additional *note fields* in the workbench. In this way the workbench tool can handle meta-data on many levels.

Finally, we are also interested in finding and designing efficient representations of abstract and dynamic processes, such as distributed decision processes, to be used for AAR as alternatives and complements to the model-based spreadsheets. We are looking into the possibility to integrate the presentation of the results into the MIND-framework, an achievement that would fulfill and close the loop from modelling of complex processes to presentation of results and performance.

References

Annett, J., Duncan, D., Stammers, R. B. & Gray, M. J. (1971). *Task Analysis*. London: Her Majesty's Stationary Office.

Bainbridge, L., Lenior, T. M. J. & Schaaf, T. W. van der, (1993). Cognitive Processes in Complex Tasks: Introduction and Discussion. *Ergonomics*, 36, 1273-1279.

Brehmer, B (2002). Självsynkronisering: En ny Form för Kraftsamling i det Nätverksbaserade Försvaret. Rapport Nr 3 från projektet NetWork Warfare. Swedish National Defence College, Sweden.

Crissey, M. J., Morin, M. & Jenvald, J. (2001). Computer-Supported Emergency Response Training: Observations from a Field Exercise. In *Proceedings of the 12th International Training and Education Conference, ITEC 2001*, Lille, France.

Dirsoft (2000). Visual Study 2000. www.dirsoft.com

Drury, C. G., Paramore, B., Van Cott, H. P., Grey, S. M. & Corlett, E. N. (1987). Task Analysis. In G. Salvendy (ed.), *Handbook of Human Factors*. New York: John Wiley & Sons.

EA Games (2003) Battlefield 1942. Software, http://www.eagames.com/official/battlefield/1942/us/home.jsp

Gasparski, W. W. (1991). Systems Approach as a Style: A Hermaneneutics of Systems. In M. C. Jackson, G. J. Mansell, R. L. Flood, R. B. Blackham & S. V. E. Probert. *Systems Thinking in Europe*. New York: Plenum Press.

Gist, M. E, Hopper, H. & Daniels, D (1998). Behavioural Simulation: Application and Potential in Management Research. Organizational Research methods, Vol 1. No 3 July 1998 pp 251-295.

Harvey, R. J. (1991). Job Analysis. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organisational Psychology 2 ed., vol. 2*. (pp. 71-163). Palo Alto, CA: Consulting Psychologists Press, Inc.

Hollnagel, E., & Woods, D. D., (1983). Cognitive Systems Engineering: New Wine in New Bottles. *International Journal of Man-Machine Studies*, 18, pp. 583-600.

Jenvald, J., Crissey, M. J., Morin, M. & Thorstensson, M. (2002). Training Novice Observers to Monitor Simulation Exercises. In *Proceedings of the 13th International Training and Education Conference, ITEC 2002*, Lille, France.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, J., Orasanu, R., Calderwood, & C. Zsambok (Eds.), *Decision Making in Action: Models and Methods*. Norwood, CT: Ablex.

Kolbe, R. H. (1991). Content analysis research: An examination of applications with directives for improving research reliability and objectivity. *Journal of Consumer Research*, 18, 243-250.

March, J. G. (1991). How Decisions Happen in Organizations. Third Annual Conference on Computer Supported Cooperative Work (1990, Los Angeles, California). *Human Computer Interaction*, 6(2), 95-117.

Morin, M. (2001). MIND - Methods and Tools for Visualization of Rescue Operations. In *Proceedings of The International Emergency Management Society's Eighth Annual Conference, TIEMS 2001*. June 19-22, Oslo, Norway.

Morin, M. (2002). *Multimedia representations of distributed tactical operations*. Linköping Studies in Science and Technology: PhD Dissertation No 771, Linköping, Sweden.

Morrison, J. E. & Meliza, L. L. (1999). Foundations of the After Action Review Process. Special Report 42. U.S. Army Research Institute: Alexandria, Virginia

Neisser, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology. San Francisco: W. H. Freeman & Company.

Rankin, W. J., Gentner, F. C. & Crissey, M. J. (1995). After Action Review and Debriefing Methods: Technique and Technology. In *Proceedings of the 17th Interservice / Industry Training Systems and Education Conference*, pp. 252–261, Albuquerque, New Mexico, USA.

Rasmussen, J. (1993). Analysis of Tasks, Activities and Work in the Field and in Laboratories. *Le Travail Humain*, 56, 133-155.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive systems engineering*. New York: Wiley.

Singleton, W. T. (ed.) (1979). The Study of Real Skills. *Compliance and Excellence* Vol. 2. Lancaster: MTP.

Strauss, A. (1988). The Articulation of Project Work: An Organisational Process. *Sociological Quarterly*, 29(2), 163-178.

Swedish Armed Forces Strategy for Research and Technology (2002). http://www.mil.se/attachments/rt_strategy.pdf. 040121

Thorstensson, M., Morin, M., & Jenvald, J. (1999). Monitoring and Visualization Support for Management of Medical Resources in Mass-Casualty Incidents. In *Proceedings of the International Emergency Management Society's Sixth Annual Conference, TIEMS 1999* (pp. 179–188), Delft, the Netherlands.

Thorstensson, M., Axelsson, A., Morin, M., & Jenvald, J. (2001a). Monitoring and Analysis of Command-Post Communication in Rescue Operations. *Safety Science*, 39(1–2), 51–60.

Thorstensson, M., Tingland, B., Björneberg, A. & Tirmén-Carelius, M. (2001b). Visualization of an Interagency Exercise in the Stockholm Underground. In *Proceedings of the International Emergency Management Society's Eighth Annual Conference, TIEMS 2001*. June 19-22, Oslo, Norway.

Thorstensson, M. & Wikberg, P. (2002). Model Based Data Collection in the Design and Analysis of Complex Military Systems. In *Proceedings of the 13th International Training and Education Conference, ITEC 2002*, pp. 25–34, Lille, France.

Waterson, P. E., Clegg, C. W. & Axtell, C. M. (1997). The Dynamics of Work Organization, Knowledge and Technology During Software Development. *International Journal of Human Computer Studies*, 46(1), 79-101.

Weber, R. P. (1990). *Basic Content Analysis*. 2nd ed. Beverly-Hills, CA: Sage.

Wikberg, P. (1997). *Using Expert Descriptions in Systems Analysis*. Umeå: Department of Psychology, Umeå University, Sweden.

Wikberg, P. (2001). Evaluation of Command and Control Systems. In *Proceedings of the 10th European Congress on Work and Organisational Psychology*, May 16-19, Prague, Czech Republic.

Wikberg, P, & Lundin. J. (2002a). Perception and Interpretation of Statistical Results in After Action Reviews. In *Proceedings of the 2002 Hawaii International Conference on Statistics and Related Fields*. June 5-9, 2002, Honolulu, Hawaii.

Wikberg, P, & Lundin. J. (2002b). *Grafiska rapporter*. Swedish Defence Research Agency, Stockholm, Sweden.

Wikberg, P. & Lundin, J. (2003). Experimental Simulation of Complex Social Systems. In *Proceedings of the 11th European Congress on Work and Organisational Psychology*. May 14-17, 2003 Lisbon, Portugal.

Wikberg, P., Persson, P-A., Holmström, H., Andersson, J. & Allberg, H. (2002) *LedSystM-spel 2: Överföring av information mellan försvarsmaktens olika ledningssystem och spel som metod för försök med ledningsmetoder.* Swedish Defence Research Agency, Linköping, Sweden.

Wikberg, P. & Thorstensson, M. (2003) Rapid Evaluation of Complex and Distributed Systems. In *Proceedings of the 11th European Congress on Work and Organisational Psychology*. May 14-17, 2003 Lisbon, Portugal.

Woods, D.D. (1993) Process-tracing Methods for the Study of Cognition Outside of the Experimental Psychology Laboratory. In G. Klein, J., Orasanu, R., Calderwood, & C. Zsambok (Eds.), *Decision Making in Action: Models and Methods*. Norwood, CT: Ablex.

Appendix 1. A model of decision processes.

Se case II (page 11) for more details.