Crisis Response Operations Theory and Practice (Do the metrics measure up?)

Mr. Russell Hayes
Technical Leader, OOTW Skill Group
Defence Scientific and Technical Laboratories
Farnborough, Hampshire
United Kingdom
e-mail:rhayes@ dera.gov.uk

George Rose, Ph.D.
Technical Manager, Front Line Department
Centre for Defence Analysis
Farnborough, Hampshire
United Kingdom
e-mail:grose@ dstl.gov.uk

Russell Hayes joined the electronics section of the Defence Operational Analysis Establishment (DOAE) at West Byfleet England in 1975 and during the next 3 years undertook the development of trials and instrumentation recording military field trials and exercises. In 1978 transferred to the Field Trials division of DOAE and conducted further analysis of exercise battles to enable reconstruction of the actions from which human factors and issues could be identified. From 1979 to 1988 Russell moved into the Land Air studies division to analyse alternative concepts of operations for the British Army in Germany. This included the development of a series of manual wargames to explore tactical and operational issues raised by novel concepts. This work led to the production of a new model that could represent innovative concepts of operations. The following two years were spent providing analytic support to the Conventional Arms Control negotiations, analysing and testing national proposals. In 1990 provided analytic support to the MoD and the Permanent Joint Headquarters during the Gulf War on campaign level issues. This was followed throughout the 90's by project managing a wide range of studies providing analysis on high-level policy issues. In 1999 Russell became the project manager for the DIAMOND model and the Technical Leader for the OOTW skill group at Defence Scientific and Technical Laboratories (DSTL).

Following four years in a steel works chemical laboratory, George Rose attended Strathclyde University, Glasgow, and gained a First Class Honours degree in Metallurgy and a Doctorate in Philosophy. His early post graduate career as a research scientist was concerned with developing novel techniques for the use of high strength non-ferrous alloys for submarines systems. Building on this experience George held research policy positions in UK Ministry of Defence and British Embassy in Washington, USA. Promoted to Divisional Superintendent in the UK Defence Operational Analysis Establishment at West Byfleet, George was responsible, in the early 1990s, for studies of major UK procurement issues in the challenging post cold war environment. Such studies as EF 2000, attack helicopter, maritime air defence, and space based surveillance. From 1994 to 1998 while with the HQ ARRC, George led a small military/civilian team developing models and staff tools for NATO's rapidly deployable HQ. This has necessitated methodology development both for

warfighting and for operations other than war(OOTW). He gave advice to senior commanders throughout the Bosnia planning carried out by the HQ and then in late 1995 deployed to Sarajevo for the year-long IFOR deployment before re-establishing the HQ in Germany in late 1996. During the time in HQ ARRC, George developed and fielded a HQ Information System Technology Demonstrator – the system was used in Bosnia, and was recognised as one of the most advance systems deployed in a HQ. Since 1998, He has been Technical Manager for the Front Line department in the UK's Centre for Defence Analysis transferring knowledge and experience of direct OA support to the military into the challenging tri-service environment. Recently He has been elected a Visiting Fellow of Cranfield University in the UK in recognition of work in information systems applications to solve military problems. In summary, George has wide first hand experience of scientific and operational analysis management and the solving of military problems in the UK and in the international field.

BACKGROUND

While measures of effectiveness are well understood for warfighting there is a wider range of potential measures available to determine the progress of Crisis Response Operations (CRO). This paper describes a comparison between one possible theoretical method of assessing the effectiveness of military forces committed to a CRO with the actual experience of analysts deployed in Bosnia and Kosovo supporting. The paper outlines the theoretical approach; the method and measures proposed, and then the technique used in theatre and finally contrasts how well the two are matched.

THEORETICAL ASSESSMENT METHOD FOR CRISIS RESPONSE OPERATIONS

THEORETICAL APPROACH

The aim was to "provide a sound foundation, based on historical records, for the development of analytical methods for assessment of Crisis Response Operations (CRO)[1]" [1]. The work was to cover the development of a taxonomy of CRO tasks that adequately described the activities undertaken during CRO. Using this taxonomy as a baseline, subsequent work would then assess the historical value of each task and identify a method assessing CRO.

A literature survey identified a series of documents [2-8] produced by the NATO Consultation, Command and Control Agency (NC3A) that contained a CRO taxonomy suitable for the project. In co-operation with the NC3A a compilation of these documents was produced to provide a suitable reference document for this project [9]. Following the identification and adoption of the NC3A taxonomy metrics were identified for the tasks in the taxonomy and an analytic framework proposed in which these metrics could inform

[1] Crisis Response Operations (CRO) is the current NATO title for what was previously referred to as Peace Support Operations (PSO). Both of these names are used synonymously with the terms Operations Other Than War (OOTW) and Non-warfighting Operations.

Measures of Effectiveness (MOE). Although it is recognised that in many cases the terms metrics and MOE are interchangeable the team attempted to apply the terms separately and uniquely to bring an artificial but expedient division between high and low level analysis.

This combination of tasks, metrics and MOE provides the structural basis for assessing the historical value of each task in CRO. Hard analytical metrics (particularly suitable for historical research) and softer MOE provide the framework under which simple analysis of CRO can be conducted. This framework, once developed, should assist us in identifying what aspects of CRO our existing tools and techniques cover and which areas require further development.

The NC3A task hierarchy was developed to identify the core tasks associated with a variety of CRO mission types TN803 [9]. Following the identification of the core tasks the NC3A team developed "Troops to Task" rules which were then applied during the NATO Defence Requirements Review 1997 (DRR97) to determine the NATO force requirements to conduct CRO. The NC3A task hierarchy is shown in Figure 1 and each of the tiers of the hierarchy are discussed subsequently.

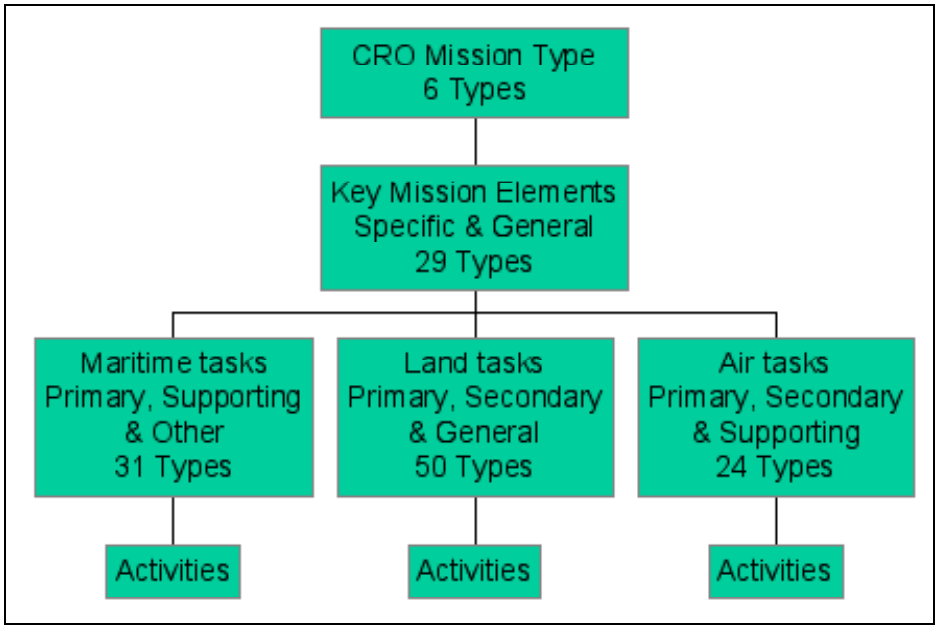

Figure 1: NC3A Taxonomy.

CRISIS RESPONSE OPERATION MISSION TYPES

The CRO mission types are Peacekeeping, Peace Enforcement, Humanitarian Aid and Disaster Relief, Peace Building, Peace Making and Conflict Prevention. For the purposes of the NC3A' taxonomy peacekeeping and peace enforcement were deemed to be the most demanding in terms of tasks and force elements. Consequently, the initial NC3A work focused on these types of operation with a partial examination of humanitarian aid included in subsequent iterations of the work.

KEY MISSION ELEMENTS

The key mission elements encompass the high level objectives of the 'military solution' to a set of generic CRO. The key mission elements were identified from analysis of the factors associated with each CRO mission type (and as stated in the previous paragraph this analysis focused on peacekeeping and peace enforcement).

The key mission elements are the high level tasks of the CRO, but are termed "elements" to avoid confusion with the lower level tasks that appear later in the NC3A taxonomy. The key mission elements themselves are of two types: specific and general. Specific key mission elements are those related to the objectives of individual operations, i.e. specific to some, but not all, peacekeeping and peace enforcement operations. General key mission elements are those which are components of all military solutions, such as protecting lines of communication. These key mission elements were developed through expert judgement and a total of 29 were identified.

TASKS

A total of 105 tasks were identified in the course of the development process. They were further sub-divided into the following categories:

- Land tasks (primary, secondary and general).
- Maritime tasks (primary, supporting and other).
- Air tasks (primary, secondary and supporting).

The key mission elements describe all the important aspects of a generic CRO mission, but they do not provide enough information to identify force requirements. Consequently all essential tasks which contribute directly or indirectly to the key mission elements have been identified. The tasks are grouped into primary, secondary and supporting tasks, dependent on whether they directly contribute to a mission element (a primary task), support a primary task (a secondary task), or provide general support to all tasks (general tasks, e.g. provide Command and Control)[2].

ACTIVITIES

The tasks themselves consist of activities, which provide an overview of what each task encompasses for the military resources allocated. Although activities are presented in the

[2] Unfortunately, the nomenclature used for primary, secondary and supporting tasks varies across the land, air and maritime domains. This work has been presented as it appears in the NC3A documents and has not been standardised for this paper. This is to provide a direct audit trail to the existing NC3A documents but does unfortunately leave some ambiguity within this paper on the correct interpretation of primary, secondary and supporting tasks.

analysis no further sub-division (or further analysis of activities) is presented beyond this point. This is not to say that activities are indivisible or cannot be described at a lower level.

LAND TASKS –- PRIMARY

This document will concentrate on the land aspects, as the apporach taken to analysis the maritime and air aspects is very similar. Primary mission tasks are defined as those that contribute directly to the key mission elements. In general, most primary mission tasks contribute to several key mission elements as is normal in military operations. Also, several tasks may contribute to the same key mission element. Examples of primary tasks include "securing seaports," "assisting host nation in refugee and displaced person control" and "controlling borders."

LAND TASKS — SECONDARY

Secondary tasks contribute directly to primary tasks by either increasing the effectiveness of the forces conducting the primary tasks (i.e. force multipliers) or by supporting the forces on the primary task (i.e. force enablers). Typical examples of secondary tasks are "providing military engineering", "providing operational reconnaissance" and "providing artillery fire support". Force components that are required for secondary task functions are usually specialised in nature. The development of the Troops to Task rules drew on the considerable military doctrine devoted to the employment of these specialised force components.

LAND TASKS — GENERAL

General supporting tasks contribute to any force component regardless of its task. The force components required for general supporting tasks, like secondary task force components, tend to be specialised in nature. Again, substantial military doctrine relevant for CRO exists for these components. Typical examples of general support tasks include "provide health services", "provide movement services for supplies" and "exercise Command and Control". Some force components normally considered as secondary (combat support for land forces) or general (combat service support for land forces) for warfighting missions were needed to fulfil primary tasks for CRO. In such cases the two aspects of the force component were dealt with separately which results in some repetition of tasks within the taxonomy.

TAXONOMY SUMMARY

The NC3A taxonomy was developed to identify the core tasks associated with peacekeeping and peace enforcement. It is structured in four hierarchical layers. The top level is the CRO mission types (e.g. peacekeeping), of which six types were identified. The next level down is the Key Mission Elements, which describe the high level mission objectives, of which 29 were identified. The third layer down are the tasks, divided into land, maritime and air, of

which a total of 105 types were identified. The final level covers activities, of which there are many hundreds, describing CRO low level actions and objectives.

METRICS DEVELOPMENT

Much of the development effort for the project was devoted to identifying metrics for peacekeeping, peace enforcement and humanitarian aid CRO. The working definition which the project used for a metric is reproduced below: *A Metric is a low-level measure of performance directly attributable to at least one task.* Each task may have more than one metric and the combination of these metrics describes the overall performance of that task. A metric should be directly measurable, verifiable and based on physical changes in a system (i.e. not a soft measure). Although this makes metrics simple to identify they do not on their own describe the progress of a CRO and relate only to the progress of the task(s) with which they are associated. Typical examples are war criminals arrested, no fly zone violations and vessels searched per day.

Although not stated in the definition, the team also tried to ensure that all metrics identified should be accessible at an acceptable cost to historical researchers and/or by the collection of operational data from current CRO activities. Subjective (or cost intensive) metrics were avoided where possible. In some cases the selected metrics were specifically written to be measured by a simple yes or no, rather than by a unit of measurement. For example, although the metric "available bandwidth versus demand" for a communications system is a useful metric it is hard to research economically (if at all) with historical analysis. In such circumstances the metric would be modified to read "bandwidth adequate? Yes or No" to simplify any subsequent historical research activity. These dimensionless metrics may be advantageous when combining dissimilar task/metric pairs[3] to quantify the rate of progress of a higher level measure of effectiveness.

In identifying these metrics the first activity was to research existing documents, and in particular the Universal Joint Task List (UJTL)[4] for existing metrics [10]. A large proportion of the metrics that were identified came from this source and are subsequently referenced against their UJTL task identity. In a few cases it was necessary to modify UJTL metrics to make them more applicable to CRO operations. Sometimes there were no suitable metrics in the UJTL and alternatives were generated and recorded by the project team. This was particularly true for the land and maritime metrics where 40% of those identified were additions or modifications to the UJTL. For the air tasks the number of new metrics required was only 20%, implying a better match between air tasks in war and CRO than in land and maritime tasks. Although many of these metrics were repeated from task to task, a total of 350 task/metric pairs were identified using approximately 150 different metrics in all. (170 task/metric pairs were identified for the land tasks, 100 for the air tasks and 80 for the maritime tasks).

[3] Task/metric pair is a unique combination, often employing metrics that appear across a number of tasks. This pairing provides a more significant context for developing MOE at a later stage. For example the metric 'own casualties sustained' has a different weighting when associated with 'training tasks' than with 'combat tasks'. MOE will be more sensitive to the task/metric pair that carries the higher political, social or operational penalty.

[4] The Universal Joint Task List is an unclassified publication issued by the US Department of Defence that provides a standardised manual describing the requirements for planning, conducting, assessing and evaluating joint and multi-national training.

In carrying out this work it became clear that metrics fell into a variety of categories. There were those that described aspects of the task or environment, (e.g. number of displaced civilians in theatre), those that described an end state (e.g. number of displaced civilians returned to their homes) and those that describe trends or progress to date (e.g. rate of return of displaced civilians). It is often a combination of all these different types of metrics that describes the task and not any one type individually. Therefore, for this work to be exploited successfully the metrics should be sub-divided into:

- Resource/environment metrics.
- End state metrics.
- Trend and progress metrics.

READING THE EXAMPLE TABLES

Figure 2 gives an example of one of the entries for the metrics table. The task is identified by both a NC3A and DERA[5] Identity number and in the second column the name of that task is given. The main activities associated with that task are listed in the third column. The final column lists the suggested metrics. Further definition, comments or simple ways of quantifying tasks are presented in [ ] brackets. Where no simple suggestion is given the metric should be described in the units that appear most appropriate to the metric, such as time, number of occurrences etc. If the metric was taken from a UJTL task then that UJTL task is referenced in italics enclosed in brackets ( ). The metrics are identified as follows:

DERA ID (NC3A ID)	**Task Descriptor**	**Activities**	**Suggested Metrics**
LAND03 (DRRG04)	Plan and Transition to Civil Administration	• provide temporary civil administration functions • provide legal services	**M1** provide temporary civil administration functions [Y/N]? **M2** Provide legal services [Y/N]? **M3** Time for normal [autonomous] civil function to be restored.

Figure 2: Example Table entry.

METRICS – SUMMARY

[5] Defence Evaluation and Research Agency soon to become Defence Science and Technology Laboratories (DSTL)

This paper has proposed that the term 'metric,' when applied to CRO analysis, should be "a low-level measure of performance, directly attributable to a task that can be measurable, verifiable and attributed to changes in a physical system." Using this definition 150 different metrics were proposed to support the NC3A task list, a large proportion of which came from the Universal Joint task List [10]. In conducting this work it became clear that metrics fell into three distinct categories: resource/environment metrics, end state metrics and trend/progress metrics. It was considered by the authors of this report that all three types were necessary to describe fully the progress of tasks in CRO.

MEASURES OF EFFECTIVENESS

INTRODUCTION

The development of a set of task/metric pairs in itself is not sufficient to describe the progress of a CRO. The metrics describe the progress of single tasks but not the aggregated affects of all those tasks working in conjunction. It is this combination of task outcomes and metrics that we refer to as measures of effectiveness (MOE). Working from this basis this chapter sets out to achieve two things. First, to define what MOE are and how they can be structured to assist us in CRO analysis. Second, to show how tasks, metrics, MOE and CRO success can be brought together in an analytical framework.

MEASURES OF EFFECTIVENESS (MOE)

The working definition for MOE used by the project team is as follows: *A Measure of Effectiveness is a high-level, aggregated measure of performance, typically describing the progress of a CRO towards meeting its goals.* Some examples of MOE include 'the level of safety provided to the civil population', 'progress towards establishing appropriate conditions for lasting security' and 'degree of protection afforded to the intervention force'. The performance of these MOE may be assessed directly from the aggregation of lower level task/metric pairs and/or through softer analytical processes that may be indirectly linked to lower level task progress. As it is likely that measuring the military mission[6] alone will not provide sufficient data with which to assess CRO performance the full MOE set must also cover the social, political and economic missions.

It is worth noting that the distinction between metrics and MOE used by the project covers a variety of issues:

a. Low level vs high level – e.g. number of patrols without incident vs regional security.
b. Measured instances vs abstract quantity – e.g. refugees repatriated vs desired end state.

[6] Military mission is defined as the sum of all tasks and activities associated with the military domain and does not mean a specific phase or operation within the CRO. Political mission refers to political tasks and activities in a similar way and so on for the social and economic missions.

c. Tangible vs intangible – e.g. murder rate vs fear of violent crime.

d. Inputs vs outputs – e.g. supply of engineering expertise vs freedom of manoeuvre.

The sting in the tail for developing and using MOE is contained within the last sentence of the definition:

> *"As it is likely that measuring the military mission alone will not provide sufficient data with which to assess CRO performance the full MOE set must also cover the social, political and economic missions."*

The OA community has on the whole limited itself to assessing MOE from the military perspective and has tended to exclude political, economic and social MOE. This is intuitive, as it is an approach that has served us well in assessing MOE for war-fighting operations. Once war begins, political, economic and social considerations often become directly tied into the results of the conflict and therefore directly tied into the military MOE. Only when the military outcome is no longer beyond doubt (or has been in doubt for too long, e.g. Vietnam) do these other factors come to prominence.

In CRO the political, economic and social mission components often have at least an equal weighting to the military mission. This is because the military operation in isolation is unlikely to deliver success, as it does not have the full spectrum of capabilities to address all the issues that are important in CRO. We must judge the success of CRO against criteria that are different from those appropriate to war-fighting operations, as even where the military component is vital, it needs to be supported by successful political and social victories for the operation to conclude successfully. As examples, divided communities must reunite, political parties must co-operate, disputes must be resolved and ethnically/politically opposed groups must tolerate one another enough for a divided country to move on. Military intervention can assist in all these issues but it cannot alone deliver any of them.

By introducing the requirement to assess MOE across the political, economic, social and military missions we increase the complexity and length of time an analytical assessment may take. We can simplify the set of MOE by introducing the concept of a unique subset, called the Key Mission Components (KMC), defined below:

- *The Key Mission Components (KMC)* are those MOE where a satisfactory performance level is *essential* for mission success. They are a subset of MOE that should always be considered first in any analysis.

Bearing this in mind it is now appropriate to define two further aspects of any individual measures of effectiveness: Threshold of Success (TOS) and Threshold of Failure (TOF).

- *The Threshold of Success (TOS)* is a recognised 'value' at which an individual MOE is considered to be succeeding or to have succeeded. This value may be derived from a combination of lower level task/metric pairs and/or soft analysis.

- *The Threshold of Failure (TOF)* is a recognised 'value' at which an individual MOE is considered to be failing or to have failed. This value may be derived from a combination of lower level task/metric pairs and/or soft analysis.

Thus, the anatomy of MOE suitable for CRO analysis is shown in Figure 3.

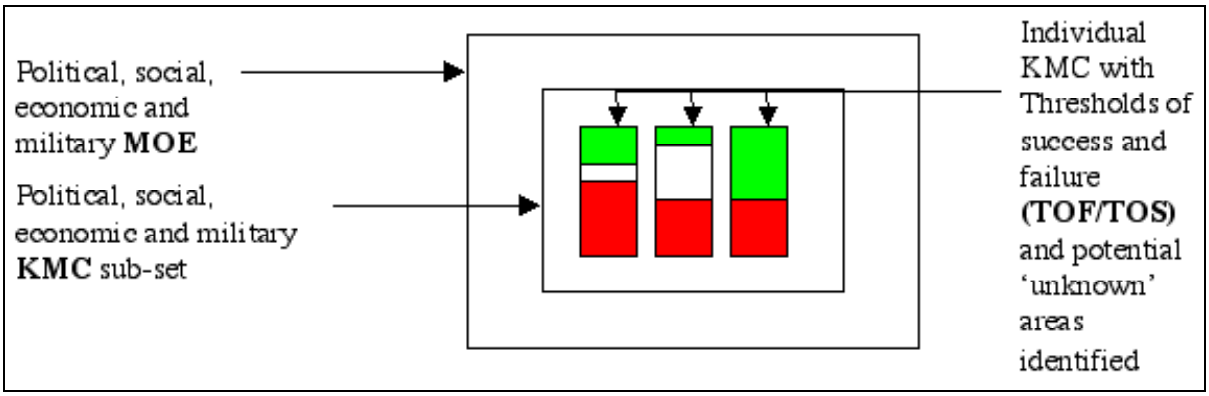

Figure 3: MOE suitable for CRO analysis.

This gives us two analysis options for CRO: we can either develop a full set of KMC and include the analysis of political, social and economic issues or we can limit our analysis to the KMC that the military can influence. If we choose the first option, and succeed, then we will be in a position to advise executive decision-makers on the likely success of future CRO. To succeed, however, we will have to develop methods for assessing political, social and economic tasks within CRO and include those within our analysis. If we choose the second option and examine only the KMC that the military can influence then we must caveat our analysis with the phrase 'the effectiveness of the military component in contributing to a successful CRO'. This is an important distinction. The success of the military component does not guarantee the success of an operation.

ANALYSIS FRAMEWORK FOR CRISIS RESPONSE OPERATIONS

To state that there is a requirement for identifying measures of effectiveness and the thresholds of success or failure at which they influence CRO is the easy part. The hard part is in carrying that through. The first stage is to identify the MOE and more specifically, the KMC. The use of historical analysis is exceptionally valuable here, as often we have declared our success criteria for CRO retrospectively once the operation is complete. We can supplement these MOE with extracts from UN resolutions, commanders' mission statements and expert judgement. There are other sources that may be tapped; for example, the NC3A analysis referenced in [9] and listed in Appendix A, Table 6, provides a selection of key mission elements which, with a little reworking would provide a basis for determining MOE for the military component of CRO.

Assuming that MOE and KMC can be identified by a dedicated attempt to draw them out of existing sources then we have most of the components we require to conduct simple analysis of CRO from high to low level. Figure 4 represents the whole system. Currently we can examine some of the aspects of the military mission at the high and low level, but not

enough to pass comment of 'the effectiveness of the military component in contributing to a successful CRO'. To do this we must develop a framework that allows us to link high and low level analysis across the full spectrum of CRO operations.

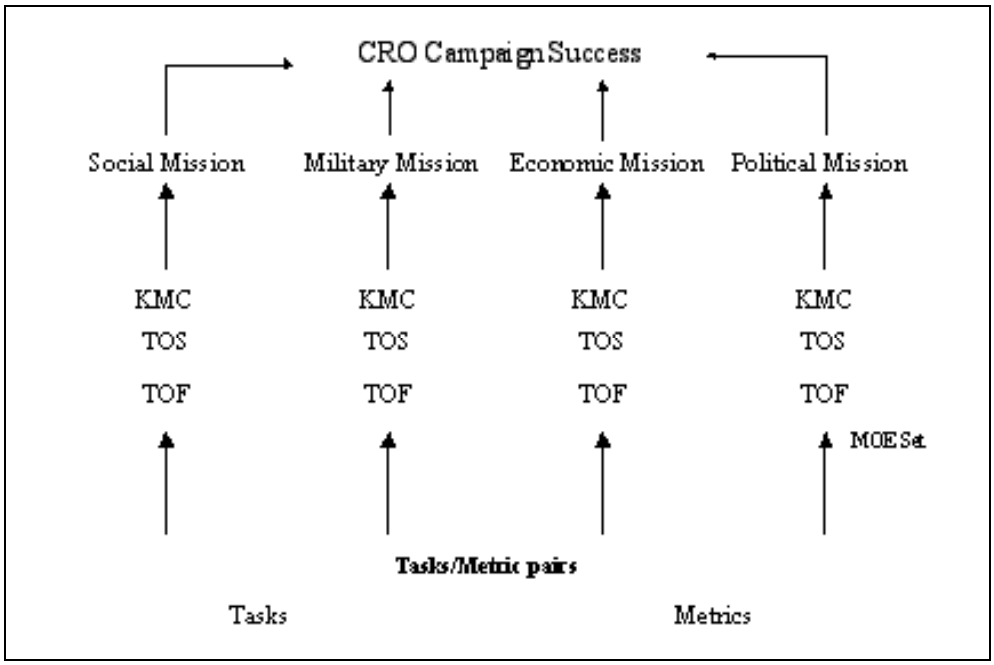

Figure 4: Potential CRO analysis structure, showing KMC channels through the MOE set.

In this system, where KMC cannot be measured directly, it may be possible to assess them by aggregating up through the hierarchical structure above using mission-oriented analysis. The important points in the context of this work are given below and shown in Figure 5.

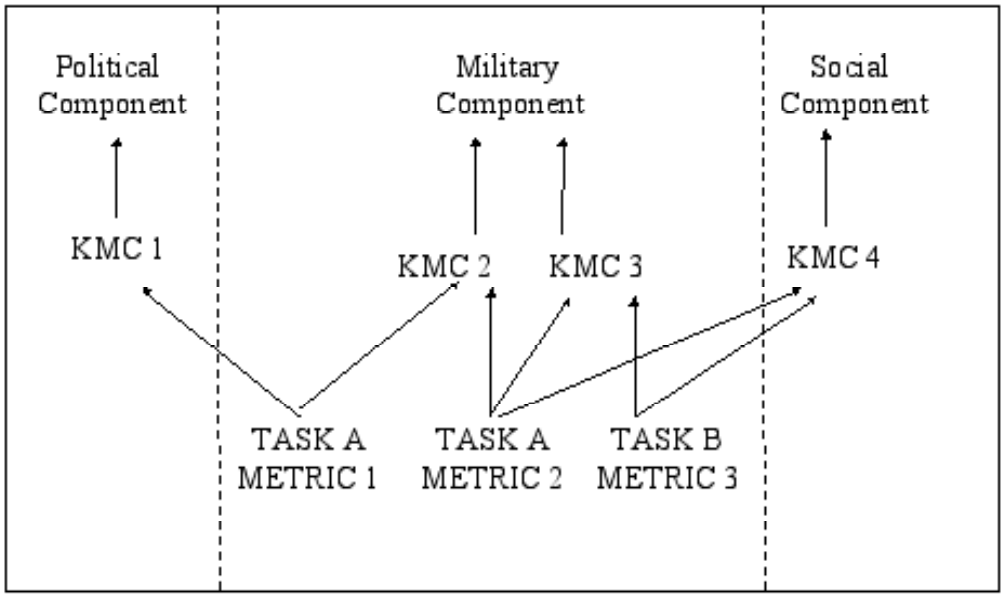

Figure 5: Mission-Orientated Analysis Example for CRO.

- Lower-level activities can support multiple tasks at higher levels.

- The logic allows alternative combinations of lower-level tasks to deliver acceptable performance at a higher level task. But, at the KMC level, satisfactory performance must be achieved in every component for mission success.

- Performance scores can be inserted at any level. If an intermediate mission component can be evaluated directly, e.g. by static scoring, then there is no need to assess lower level tasks that support only that component.

- A three level, 'traffic-light,' performance indicator makes it easy to identify affluence as well as shortfall.

In identifying TOS and TOF we should first build from the bottom up by constructing a generic list of necessary lower-level military tasks and the metrics by which they may be measured. If (scenario specific) target levels of activity are set[7], then these can be used as measures of performance for the force component(s) executing the tasks. Performance in the lower level tasks normally supports higher-level mission components, either directly or through intermediate-level aggregation.

In addition to the bottom-up determination of KMC, performance from lower level tasks and metrics can be combined with a top-down perspective. This approach, common to mission-orientated analysis, accepts that while success of the military mission is often necessary for a successful CRO, it is not sufficient. CRO are about inducing a desired behavioural change in state or non-state actors, under heavy constraints on the peace support forces' use of violence. Additionally, the initiation and termination of CRO are essentially (and intensely) political processes. The military contributions can then be thought of as: Managing[8] the conflict situation for long enough that the Strategic Information Operation (SIO) campaign can achieve the desired behavioural change(s). Conducting 'operations for strategic effect' as required in support of the above campaign.

Adding the top-down perspective has the benefit that it helps distinguish activities that contribute to strategic as well as operational success. And those that are essentially self-generated by the fact that a particular choice of force element has been deployed. Within the mission-orientated analysis approach this raises the implication that we can investigate two potential failure modes for the military component of a CRO.

- Failing to achieve the military mission (i.e. the military KMC) of a CRO.

- Undermining one or more of the non-military components of a CRO.

On the latter point that mission-orientated analysis gives us an additional perspective. For example, to ensure security of the civil population the military mission could enforce a curfew. This would however undermine other KMC, such as freedom of movement or societal reconstruction. For analysis purposes, it is helpful to have these KMC drawn out

[7] Which may be zero if the task is not required in a particular scenario.

[8] Managing, including through actual or threatened tactical use of force, the level and type of violence occurring in a region.

explicitly rather than embedded as conditions (often implicit) on the military activity. Where explicitly identified, rather than dropping an activity from the range of allowable responses, it may be possible to mitigate the adverse effects by requiring additional activities or capabilities – not all of which may be in the military sphere. Identifying these failure modes also allows an assessment of the responsibilities assigned to the forces in comparison with their capabilities. Additionally, the route by which KMC lead to failure may be different from the way in which success measures aggregate up, and this needs to be identified.

The top-down and bottom-up perspectives can be reconciled and combined, as shown in Figure 6. The degree to which this whole edifice needs to be populated with quantified data in order to inform decision-making depends on the issues to be addressed; the criterion is fitness for purpose. For example, to discriminate between alternative force element or equipment options, it may be sufficient simply to compare their cost-effectiveness in meeting a single KMC, since by definition this is mission-critical. Even in this case, however, there should be a check that all options are at least adequate for any other KMCs that they support, and that they do not introduce unacceptable failure modes – hence the KMC-TOF analysis.

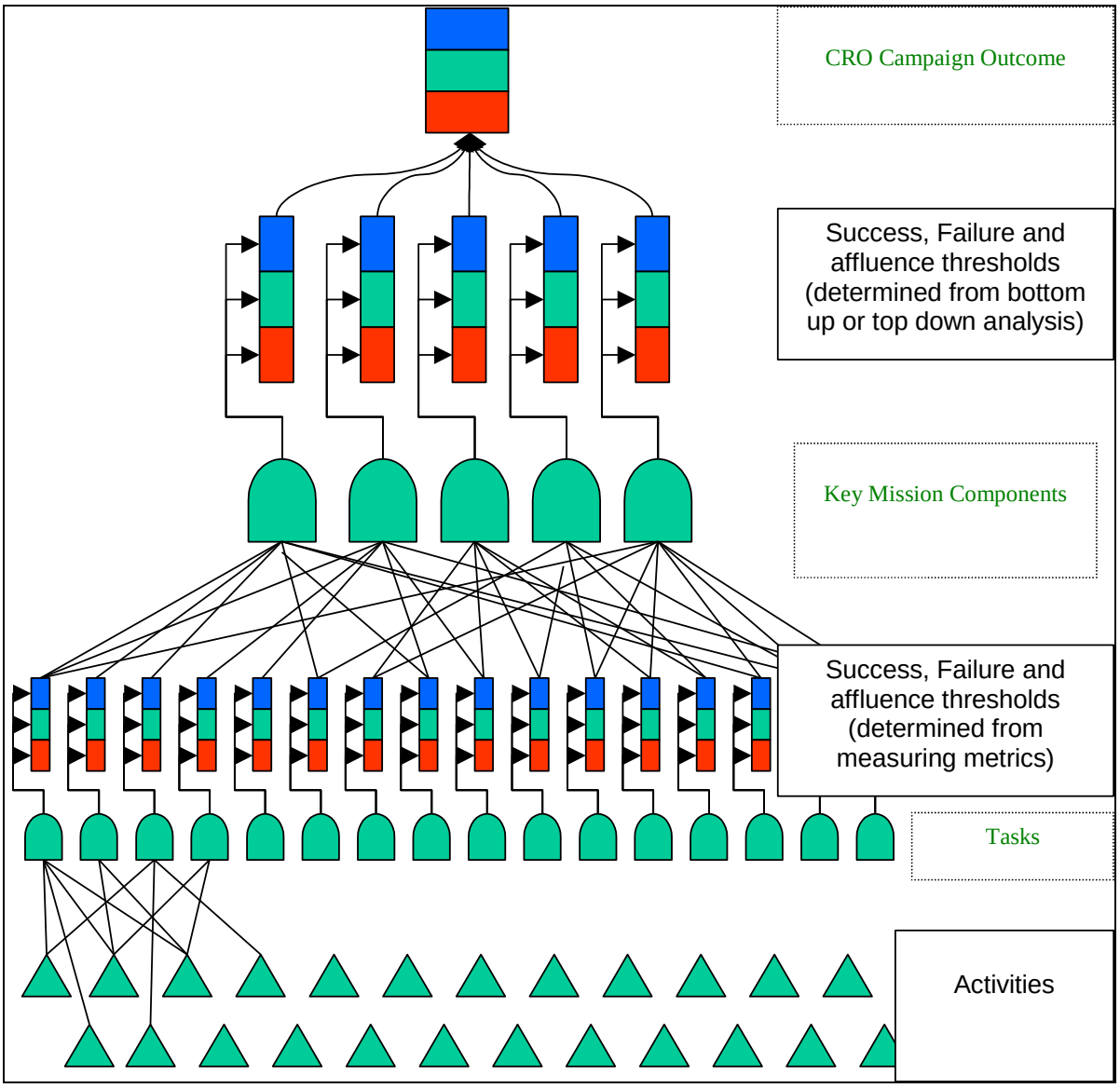

Figure 6: The mission-orientated analysis framework for CRO.

For operational planning or politico-military gaming the 'vertical slice' described above is inappropriate. Here breadth of coverage is required, but only as far down as necessary. In some circumstances, it may be permissible to mix quantitative analysis and qualitative judgements in the same structure – at least they are then visible and open to challenge.

In the case of requirements definition[9] it is important to cover a wide range of scenarios. This is particularly true in CRO, where force packages are generally small and individual force elements or equipment may have to contribute (significantly) to more and different activities in each scenario. The breadth versus depth trade-off will be on a case by case basis and is likely to lie between the more clear-cut examples above. If the lower-level performance is not too scenario specific then mission-orientated analysis can be used to analyse the whole portfolio of scenarios at once. This is particularly useful for exploring 'what-ifs' and the impact of substituting one force element for another.

GREATER UNDERSTANDING OF SUCCESS CRITERIA FOR CRO

By developing and testing the mission-orientated analysis framework described above the relationship between task success/failure, and higher level performance (at the MOE/KMC level) could be investigated. This would allow us to use task-based metrics to generate the results of some MOE/KMC and combined with a top-down assessment of other MOE/KMC it may be possible to evaluate the performance of the military mission during a CRO. Even where our knowledge of thresholds for success is insufficient to guage the progress of a CRO it may be sufficient to show that the military mission does not jeopardise other aspects of the mission (political, social etc.) through examination of the thresholds of failure on KMC.

MEASURES OF EFFECTIVENESS (MOE) FOR CRISIS RESPONSE OPERATIONS

INTRODUCTION

The UK has a long and distinguished past in its use of Operational Analysis (OA) to support military operations. The term Operational Analysis itself dates from the Second World War. More recently land and air analysts deployed to theatre to support UK forces during the Gulf operations, at the same time as analysts were supporting planning staffs in United Kingdom.

More recently the experiences of UK OA support to recent Balkans operations gives us an opportunity to assess whether operational analysis techniques have been able to contribute to the overall success of the military mission. A series of general characteristics of Measures of Success/Measures of Effectiveness were derived from observations on the work done to support IFOR, by Rose and N. Lambert (most recently reported by Lambert in a paper presented at the EURO XV11 Conference on OR, Budapest, Hungary, July 2000). This work concluded that in peace support operations it was essential that any metric used to identify 'mission success or mission effectiveness' should exhibit the following aspects:

[9] 'Needs & numbers.'

- Mission Related. They should not just be based solely on the few specific military tasks assigned to the commander but encompass the wider objectives (political, economic, social etc) of the operation.

- Comprehensive. No single measure can be expected to capture all of the information relevant to the mission. Thus any set of measures should cover a range of aspects of the mission and change in line with mission developments. It is important not to focus solely on the security aspect of the mission but include metrics that covers the wider objectives of the operation.

- Meaningful. Measures must provide a meaningful metric to progress. Those associated with tasks should focus on the effectiveness of the tasks and not simply on the accomplishment of the tasks.

- Measurable. Indicators must be able to be measured consistently and accurately over time and space (in order to see trends and compare differences). Although measurable at the beginning of a mission, the availability of the data may however prove to be transitory. As it is difficult to be prescriptive at the outset of an operation some redundancy in the data set should be planned for and the analysts should be flexible and innovative enough to use only the data that is deemed to be reliable.

- Sensitive. As the purpose of the measures is to observe trends (showing progress or no progress) measures must have sufficient resolution to enable analysts to detect that the situation is changing. The measures must also be responsive enough to show changes within the timescale of the mission.

- Timely. The measures should be responsive enough to the changes they are trying to measure so that the commander can detect the changes and act on them.

- Cost-effective. The number and types of measures used to assess effectiveness should not be a burden on those collecting and measuring them. This must be balanced against the potential benefit derived. The use of troops to collect data here is an important consideration. However, if surveys can be tied to another operational requirement, such as the creation of a perception of presence (as in the IFOR operation) this considerably reduces the burden on the troops.

- Culturally and Locally Relevant. The local customs, food, social structure, economics etc of the country that the intervention force in deployed in will be different to the countries from which the troops (and analysts) have been drawn from. Detailed work on the metrics must be carried out in theatre with local experts. Controls and benchmarks if used must be based on local standards.

The analysis methodology which has been used was originally developed by Lt. Col. John Musser US Army, at a workshop in Canada in 1999. He suggested that in support of smaller scale contingency operations there is a definite shift away from traditional combat modelling and its associated Measures of Effectiveness. What is needed is a wider approach

to Measures of Effectiveness, including new metrics when necessary to describe the data that are collected/observed during these new operations. If a MOE is to be used it should answer the following questions:

- What is the problem?
- What resources are available?
- What progress is being made?
- What are we (ultimately) attempting to achieve?

Musser suggested the following descriptors be used:

- *Resource MOE:* Numbers of troops in military units, levels of supplies, equipment holdings etc.
- *End-State MOE:* Measures the degree to which end state is accomplished. These are closely aligned with exit criteria.
- *Situational or Trend Derived MOE:* These criteria take the 'pulse' of aspects of the society in which the operations is taking place. This could incorporate assessments of the reestablishment of society, economy etc and can give general indications of the 'health' of the region of operation. These indicators can play a part in assessing the 'stability' of the area of operations.
- *Scenario Descriptors:* These are quantitative measures used to clarify/classify demographic features of the region. For example - numbers of evacuees, ethnic balance, freedom of movement, crime figures.

It is interesting to see whether recent analysis can support Musser's paradigm. The authors have reviewed the NATO and UK analysis given to supporting forces in Bosnia-Hercegovina and Kosovo and catalogued them according to the Musser Measures of Effectiveness. Examples of the types of analysis are given.

RESOURCE MEASURES OF EFFECTIVENESS

The following list of analyses projects have been carried out in support of these MOEs

- Wargaming contingency plans (BH and K), HQ information flow and structure studies (BH and K), (re)deployment modelling (BH), AWC/MWC support (K).
- Optimum positioning of reserves for election support(K).
- Infantry reserves (K).
- Troop deployment times (K).

- Transit times between towns (K).
- Surveillance unit requirements (K).
- Questionnaire analysis and advice for Information Operations (K).
- Database design advice to HQ Branches (BH and K).
- HQ Information Flow analysis (BH).
- Measurement of combat capability (BH).
- Intelligence Preparation of the Battlefield- questionnaire preparation (BH).

Examples in this category include wargaming contingency plans to investigate potential land force options and force characteristics (Figure 7). At a lower level recent support in both Bosnia and Kosovo investigated the most cost-effective use of resources within UK headquarters, particularly the efficient flow of information. Also in this category is (re) deployment modelling in Bosnia to support IFOR (Figure 8), where the analysis was looking at the implications of different force redeployment options. Also, much of the work carried out by the UK Air Warfare Centre in support of the Kosovo air campaign was concerned with the appropriate use of resources- aircraft, weapons, doctrine and concept of operations.

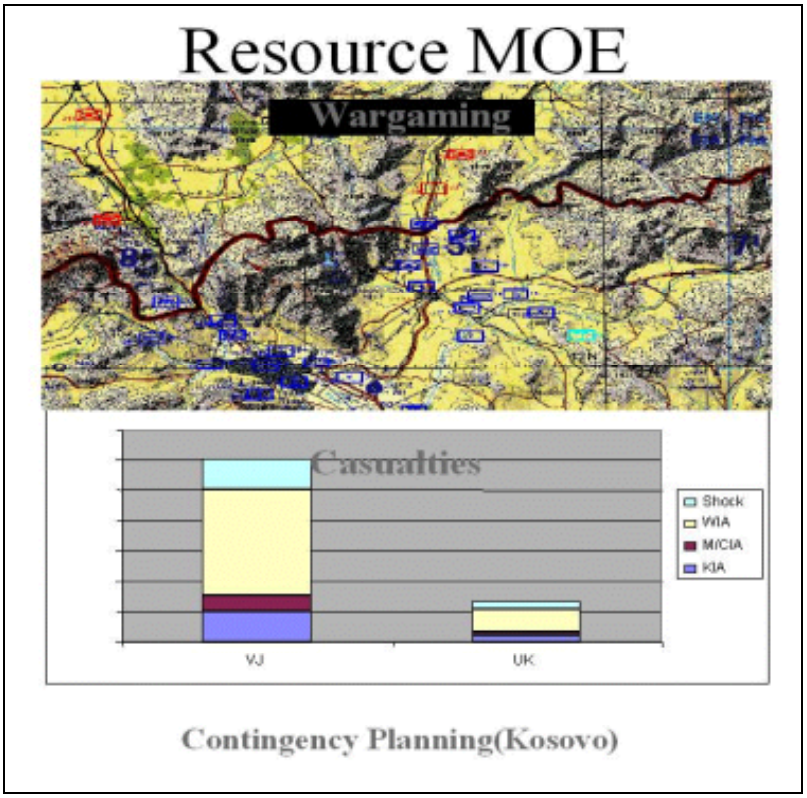

Figure 7: Resource Measures of Effectiveness.

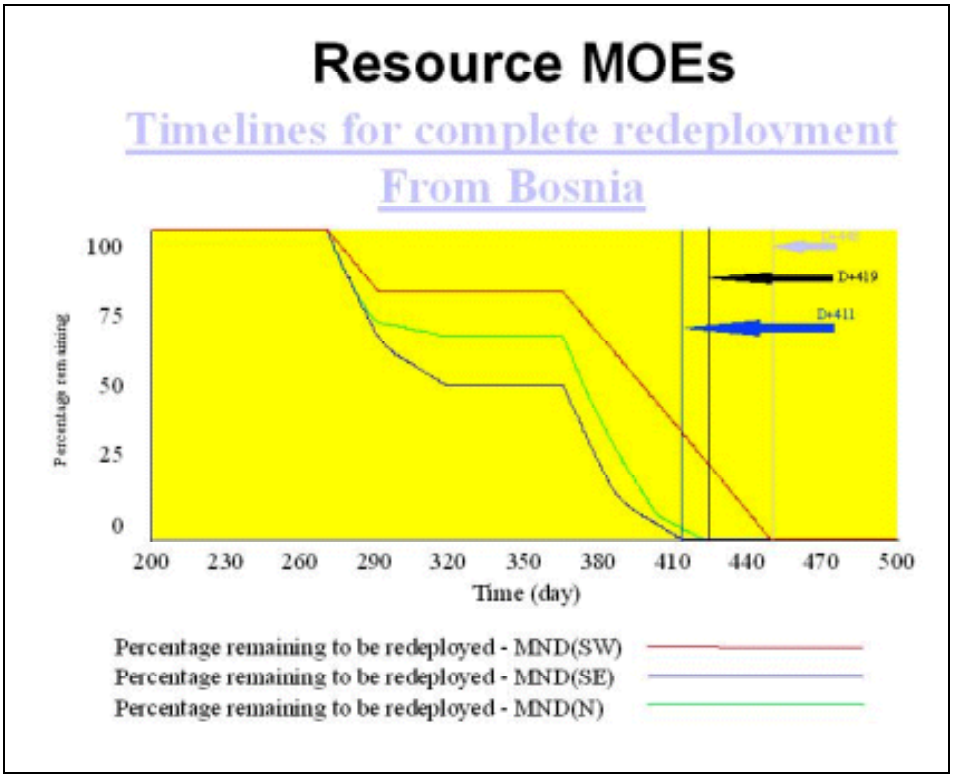

Figure 8: Resource Measures of Effectiveness.

END-STATE MEASURES OF EFFECTIVENESS

The following list of analyses projects have been carried out in support of these Measures of Effectiveness.

- Compliance of military factions (BH) / KLA to agreements (K).
- Refugee/displaced persons return modelling (BH and K).
- Election support (BH).
- Population flow – use of bus and train services (K).
- Prosperity- shop prices, ranges of goods, towns versus countryside (K).
- Popular attitudes towards ethnic minorities - changes with time (K).

A number of examples of End-State MOEs have been used. For example the compliance of various factions to signed agreements has featured in both Bosnia (Dayton Peace Agreement) and Kosovo (Military Technical Annex) Figures 9 and 10 show some typical analysis. Also a major feature towards an internationally accepted end state during IFOR was the support given to the national elections in September 1996. Simple systems dynamics modelling was a powerful tool in assessing a range of aspects in the election process.

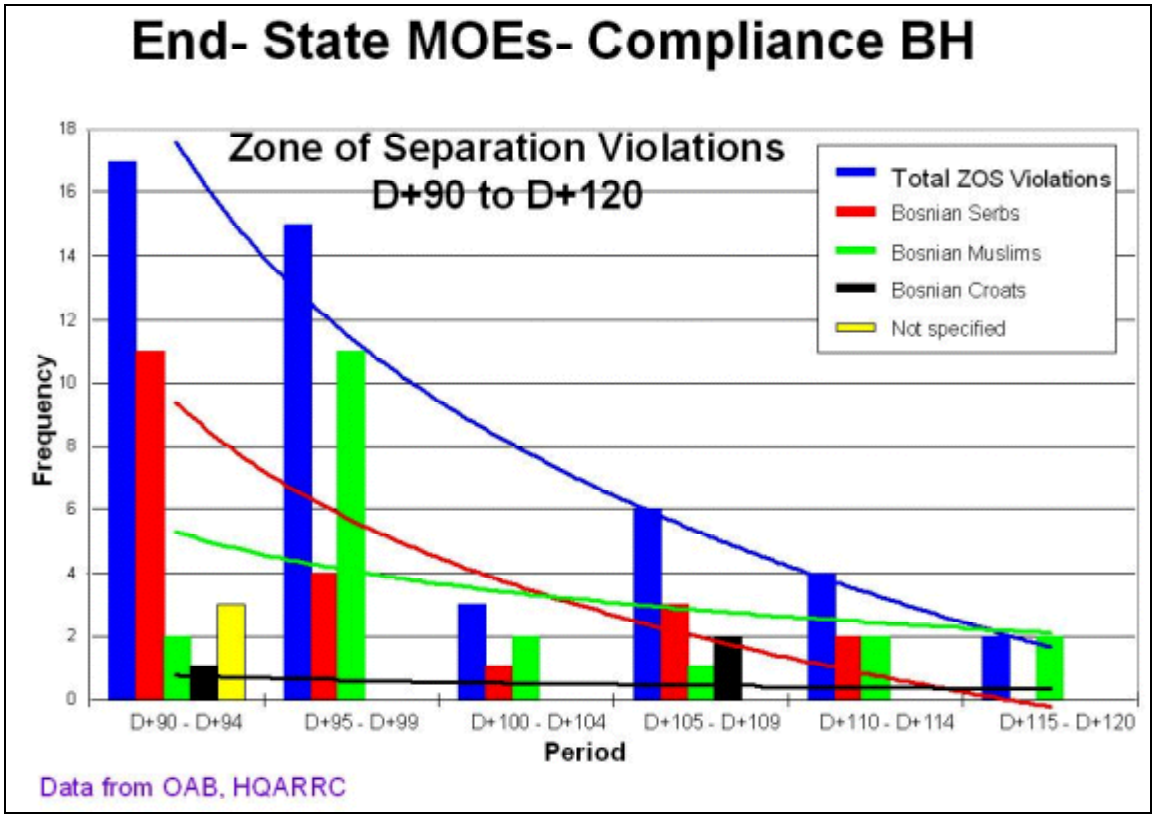

Figure 9: End-State Measures of Effectiveness.

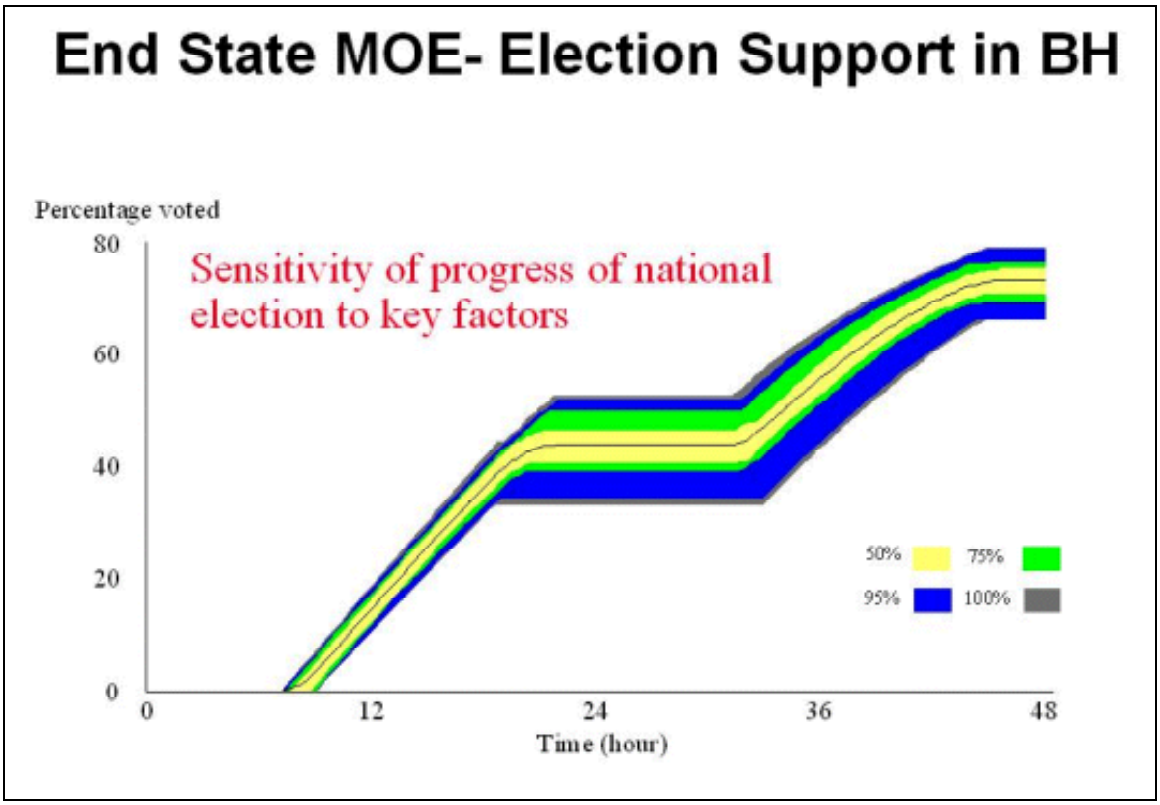

Figure 10: End-State Measures of Effectiveness.

TREND/SITUATIONAL MEASURES OF EFFECTIVENESS

The following list of analyses projects have been carried out in support of these MOEs

- Rapid Village Assessment (K), Traffic Studies (BH), Legal Infrastructure Survey (BH), Normality, SFOR/NC3A '6 Month Reviews'(BH).
- Potential reasons for non-return of Albanian IDPs (K).
- Impacts on closing down Mitrovica lead smelting works (K).
- SFOR publications survey (BH).

Both Bosnia and Kosovo have shown that some changes only occur slowly in the longer term and hence a series of Measures of Effectiveness, which can be tracked to demonstrate change, need to be employed. The 'Normality' work during IFOR is one such example, whereby collection of data in a number of simple categories allowed a timely review of changes throughout the country. In the Kosovo campaign KFOR and the UN pooled their data resources to allow a range of metrics to be identified. Figure 11 shows the availability of key grocery supplies in Bosnia. Figure 12 shows how much damage was observed to towns/opstinas throughout the country and it gave clear indications where resources/aid could be most effectively channelled.

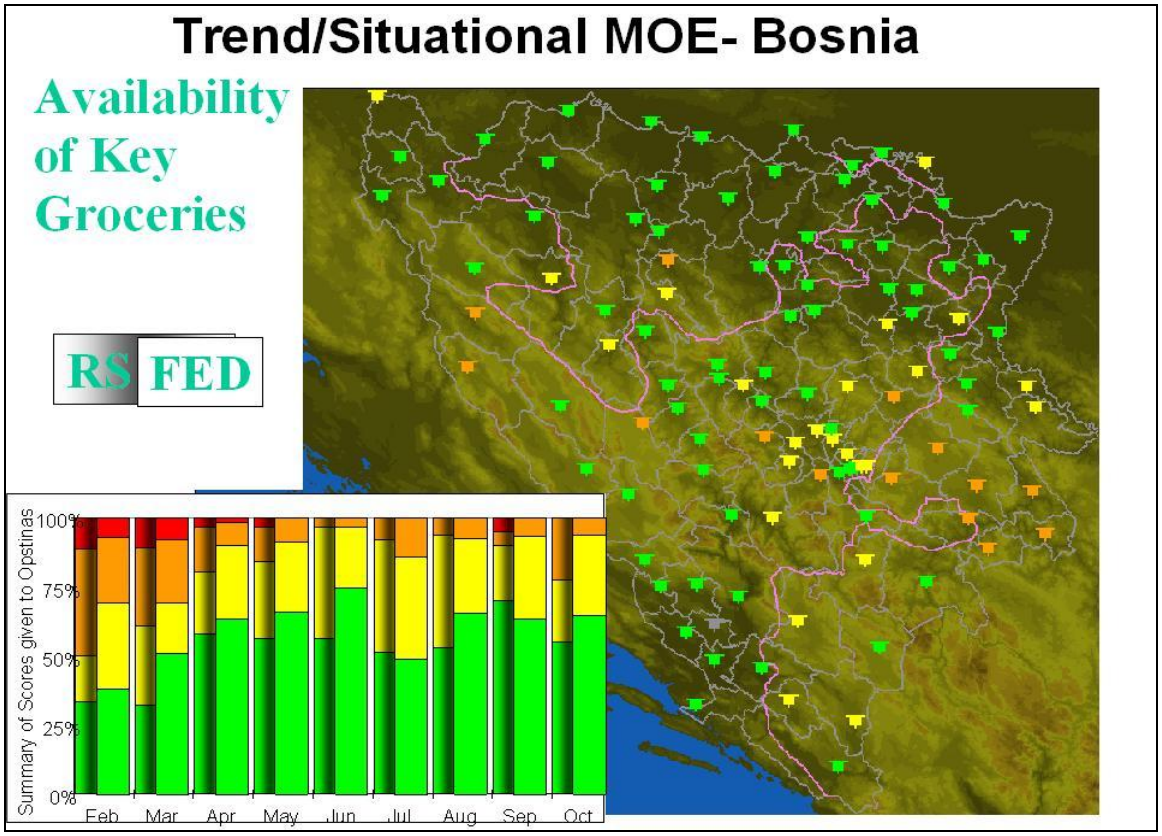

Figure 11: Trend/Situational Measures of Effectiveness.

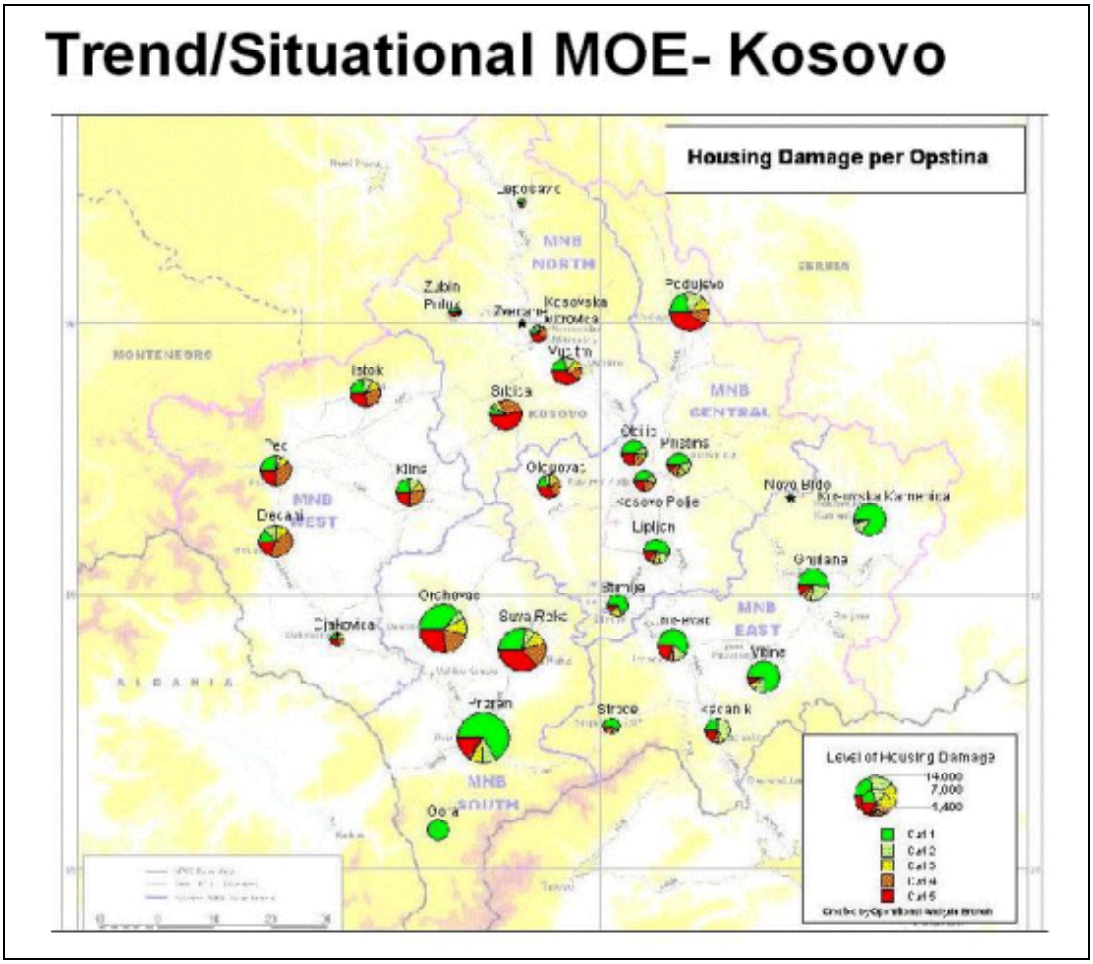

Figure 12: Trend/Situational Measures of Effectiveness.

SCENARIO DESCRIPTOR MEASURES OF EFFECTIVENESS

The following list of analyses projects was carried out to support these measures of effectiveness.

- Ethnic balance (K), freedom of movement (BH).
- Crime statistics (K).
- Prediction of DPRE returns (BH), DPRE tracking database (BH).

To aid the general understanding in an area of operations, wider descriptors are often necessary. For example in Bosnia, studies of traffic flow have been a feature of analysis. This analysis can help in the understanding of freedom of movement issues and the gradual reestablishment of commercial infrastructure. Figure 12 shows the changing pattern of traffic flow in Bosnia. Figure13 shows some 1997 SFOR analysis One key feature of the ground deployment of NATO troops into Kosovo was their impact on the number of instances of criminal behaviour (Figure 14). Collating statistics from the Military Police allowed the KFOR commanders an insight into how well the troops were re-establishing the rule of law.

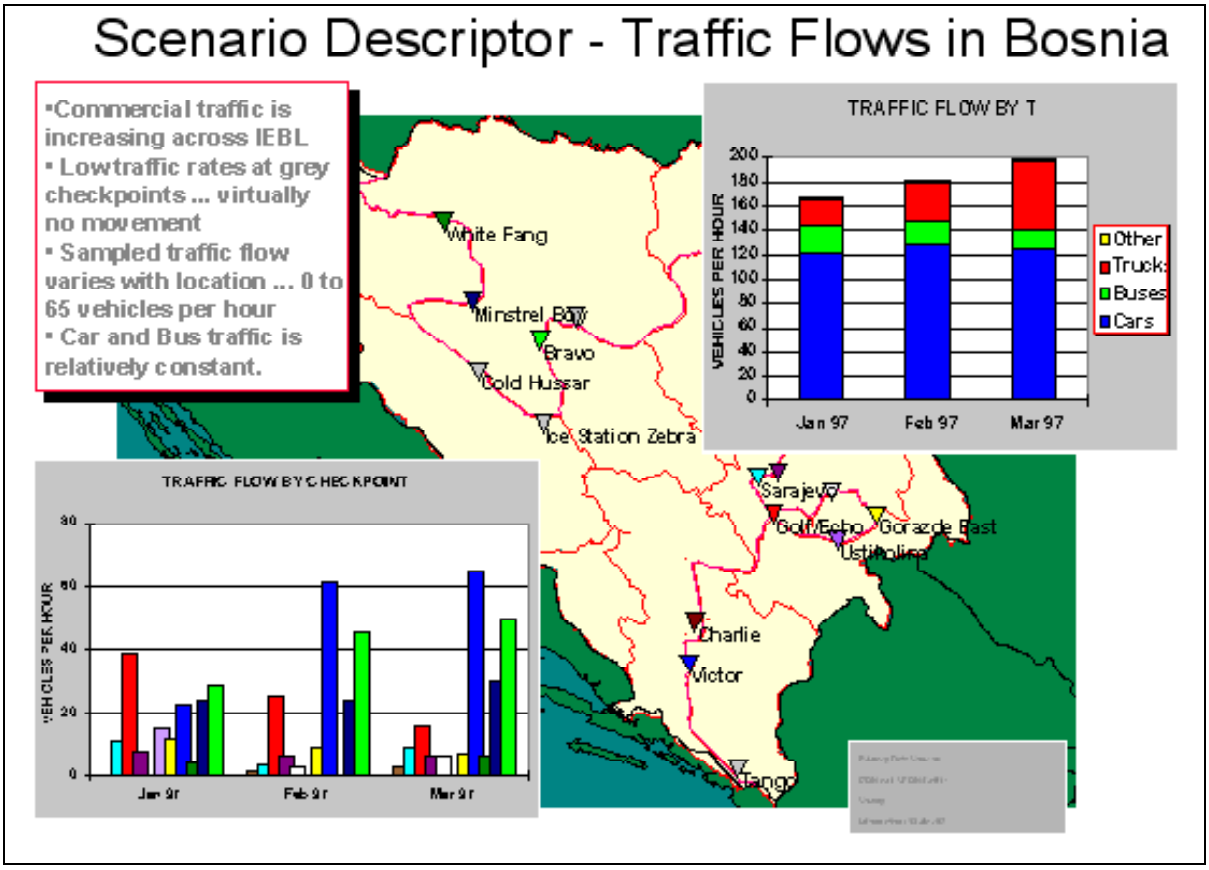

Figure13: Scenario Descriptor Measures of Effectiveness.

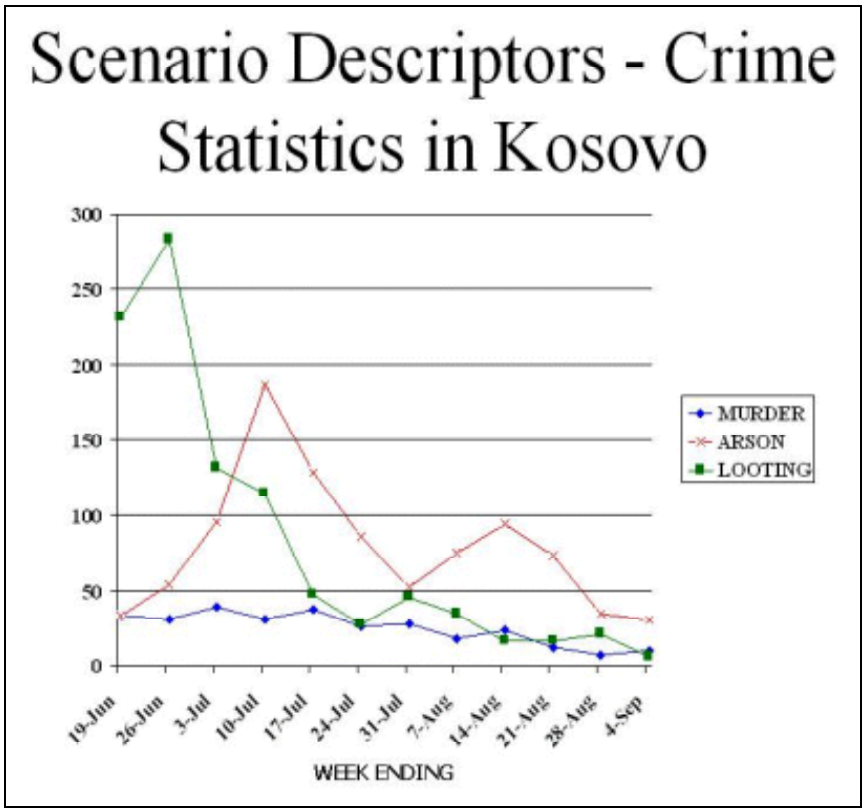

Figure 14: Scenario Descriptor Measures of Effectiveness – Kosovo crime statistics.

COMMENTS ON MEASURES OF EFFECTIVENESS

The above simple collation of analyses demonstrates that operational analysts have been able to give support in the difficult area of defining Measures of Effectiveness or success. This is early days for these analytical techniques and no single approach is yet available, although a range of topics has successfully been supported. More research is needed into Measures of Effectiveness in Peace Support Operations, especially the collection of relevant data from ongoing operations.

COMPARISON OF LAND METRICS AND DATA COLLECTED ON OPERATIONS

An assessment has been carried out to see if the following questions can be answered:

- Are data available from Bosnia/Kosovo to allow quantification of the Land Metrics developed in the first section of this Paper?
- Which metrics had data, and had these data been used for in theatre analysis?
- Where did the data come from for 'in theatre' analysis?
- What can we learn from Bosnia and Kosovo analysis support?

In order to analyse the metrics, a simple Microsoft Access database was developed. This took the list of tasks, activities and metrics already developed and eliminated the duplicate metrics such that 127 remained. Each metric then had four factors assigned to it:

1) The agency responsible for the collection of the source data (both primary and secondary).

2) The collectability of the data, scored as:

   a) Has been collected.
   b) Could be collected.
   c) Never can be collected.

3) The ease of collection of the data, scored as:

   a) Easy to collect once identified.
   b) Requires time OR effort.
   c) Requires time AND effort.

4) The time of collection of the data, scored as:

   a) During the course of the operation.
   b) After the operation has ended.

Additionally a note was made of whether the source data for the metric was collected (and if so, in what format) in both Bosnia-Herzegovina and Kosovo and whether it was used for the OA done in-theatre. This analysis and scoring was undertaking by analysts with experience of the IFOR and KFOR operations.

After all of the metrics were scored, the results were analysed at both the metric and the task level. Only the analysis at the metric level is presented here.

On the positive side, we concluded that only 10 of the 127 metrics were dependant upon source data which either could not be collected or for which the level of time and effort required was considered to be too great (Figure 15). Of the remaining 117, 69 were believed to easy to collect and a further 15 could be collected with a relatively small level of effort (i.e. 66% of all the metrics). From our experience of operations in the Balkans we believe that 49 metrics have supporting data from both the IFOR and KFOR operations. A further three have data from IFOR only, 7 from KFOR only and 30 were deemed to be irrelevant to those operations (although they may be relevant to other PSO). This leaves 28 metrics for which supporting data was not collected but could have been if either the effort or the need had been identified.

On a slightly negative side, we then did an analysis of the work that was undertaken by the OA teams in support of IFOR, SFOR and KFOR (Figure 16). Only 18 of the metrics could be supported by the data used for these in-theatre studies. Of these, 10 are covered by the Compliance work alone. The reason for this is that the metrics only describe the military aspects of the mission, in other words, the military management of the campaign. However, they do not allow an analysis of the overall, wider mission (i.e. taking account of the Political, Social and Economic aspects).

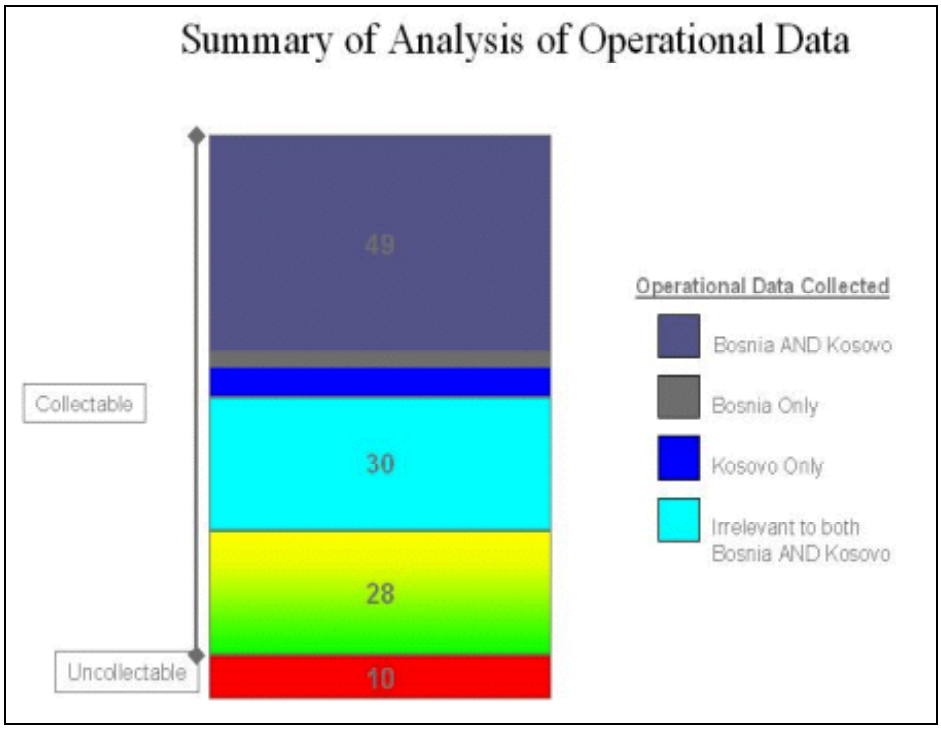

Figure 15: Summary of Analysis of Operational Data.

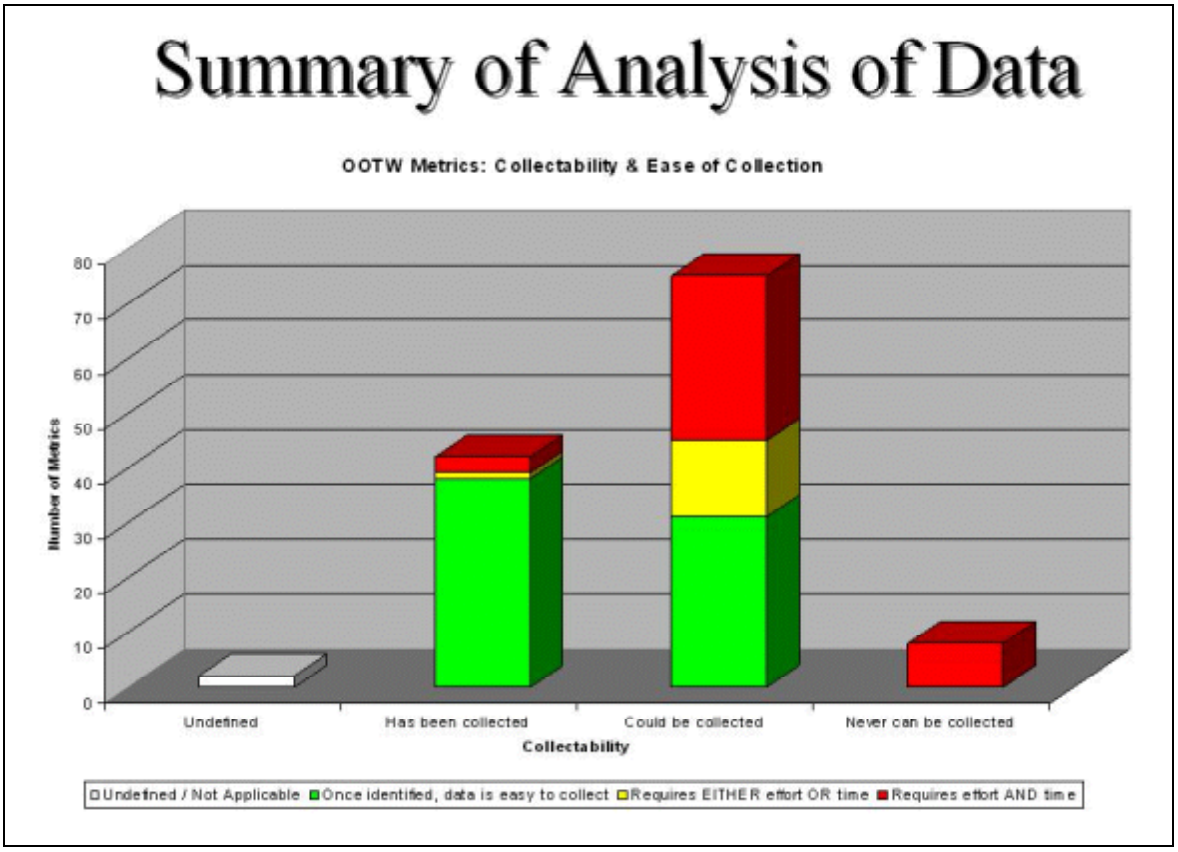

Figure 16: Summary of Analysis Data.

OA ADVICE FROM IN-THEATRE OA TEAMS

When reviewing the OA advice given to IFOR, SFOR and KFOR, it has been sought when military 'rules of thumb'/ experience is incomplete or there has been significant changes to military conditions eg, wargaming options, direct support to HQ Branches, information management. In addition, civilian aspects of the mission have been analysed and it has been shown that data has not been available through military reporting chain. In these instances fairly simple data analysis techniques required (complex models are not necessary).

CONCLUSIONS

- *Military Metrics:*
  - Lots of data collected which could be useful for assessment.
  - Data used almost exclusively for Compliance and direct HQ support.
- *Operational Analysis:*
  - Major contributions on wider civilian aspects of mission.
  - It uses data the military are not responsible for collecting.

- Non Military Metrics:
  - Not yet defined comprehensively.
  - We need to impress on Commanders the wider aspects of the mission.
- Modelling of CRO/PSO:
  - Needs Military and civilian metrics to be comprehensive/robust.

OVERALL CONCLUSIONS

1. A set of theoretical metrics can be defined for the military aspects of a CRO and operational data can be collected allows quantification of these metrics.
2. Data to assess the wider aspects of a CRO is not normally collected by the military information systems and links with other organisations need to be established.

ACKNOWLEGMENTS

Our thanks to analysts at HQ ARRC, NC3A and CDA for allowing us to use their analysis as examples of recent support given to operational commanders. The opinions in this paper are solely those of the authors and do not necessarily constitute those of CDA, DERA, UK MOD or NATO.

REFERENCES

1. Technology Group 11 Operational Analysis, Programme document, DERA/CDA/HLS/CR990005/1.0, March 1999, UK RESTRICTED.
2. Chouinard, P., 1999. A methodology for the derivation of force requirements for peace support operations, Technical Note 707, Published by NATO NC3A, March 1999, NATO UNCLASSIFIED.
3. Validation and Harmonisation of NC3A PSO Defence Planning Methodology, Final report, CR1139/TR1, Published by CORDA Ltd, October 1999, NATO UNCLASSIFIED.
4. Tasks and Rules Explained, Technical Note (to be announced), Final Draft by CORDA and NATO NC3A, 19th October 1999, NATO UNCLASSIFIED.
5. Fisher, M.J., Maritime Tasks, Unpublished, NATO NC3A, NATO UNCLASSIFIED.
6. Derivation of Requirements for PSO (Air Tasks), Unpublished, NATO NC3A, NATO UNCLASSIFIED.
7. Chouinard, P., Peacekeeping Mission Tasks, Unpublished, NATO NC3A, NATO UNCLASSIFIED.
8. Chouinard, P., Peace Enforcement Mission Tasks, Unpublished, NATO NC3A, NATO UNCLASSIFIED.

9. Chouinard, P., M. Fisher, M. Crookenden and A. Clements, 2000. Compilation of NC3A task taxonomy and troops to task Work (Draft), TN803 June 2000, NATO UNCLASSIFIED.

10. The Universal Joint Task List, Version 4.0, CJCSM 3500.04B, October 1999, UNCLASSIFIED.

11. Coyle, J.M., D. Exelby and J. Holt, 1999. Systems dynamics in defence analysis: some case studies, *Journal of the Operational Research Society,* April 1999.