Mission Assessment in Complex Operations: Canadian Lessons from Afghanistan

Peter Dobias

The assessment of mission effectiveness has been called the "Achilles' heel" of the effects based approach to operations. This is especially true in the multi-agency, multi-stakeholder environment of present-day counterinsurgency. This paper addresses some of the pitfalls of assessment and suggests possible solutions to enable effective mission assessment in a complex environment. It is based to a large degree on the author's experience with the whole-of-government environment gained while supporting the mission assessment for the Canadian mission in Afghanistan both in Canada and while deployed with Task Force Kandahar. The key point of the paper is that great care has to be taken when using lower-level assessments to obtain higher-level situational picture for the state of insurgency and/or economic and governance development. Otherwise, a potentially skewed picture can arise, or the assessment becomes too laborious and ineffective.

Introduction

It is almost generally accepted that the conflict in Afghanistan features many of the characteristics of a complex environment. There are multiple actors (social and political groups) with a variety of interests, for example different ethnicities, tribes and tribal groups, the international community represented by the International Stabilization Afghanistan Force (ISAF), the United Nations Mission in Afghanistan (UNAMA), non-government organizations, Taliban and other insurgent groups, criminal gangs, and warlords. These stakeholders interact with each other in multiple ways (economically, militarily, and politically) and there are overlaps between them in time and space (e.g. various social groups can span a geographical region, or share common or contradictory interests at a point in time). Consequently, any change in one of the entities can have varying effects on the entire system. These effects can be further reduced or increased by the changes in other entities caused by the original change. For example, advances in governance can decrease the support for insurgents, but can lead to a pushback from criminal elements and warlords that perceive the advances as a threat to their power. An increase in crime-related violence might lead in turn to a decrease in support for the government and increased support for the insurgents.

The dynamics happen on a variety of scales in time, space and purpose (considering the interests of various stakeholders). Spatially, the dynamics at the village or community level are different from those at provincial or tribal/ethnic area level, and different still from the national and regional dynamics. However, these levels interact via human actors that often span multiple

levels. For example, President Karzai represents Afghanistan internationally, while at the same time he is the representative of the executive power at the national level, and he appoints governors thus impacting the provincial level as well. In the temporal realm, the security situation typically varies on a short-term basis (weeks, months), while governance and development is usually longer term (years, decades). But these different aspects of the mission affect each other. The non-linearity and complexity of the environment in Afghanistan has another impact on assessment. It is embodied in the fact that the counterinsurgency environment cannot be decomposed into small independent parts, and consequently, the constructs such as averages have a very limited meaning. Therefore, constructing a higher-level assessment as an average of the lower-level assessments can lead to significantly skewed and potentially very misleading results (see below).

Such complex dynamics presents a great challenge for mission planning and assessment of the desired progress [1,2]. Assessment must in essence answer two key questions: 'Are we doing things right?", and more importantly "Are we doing the right thing?" While the first question addresses the issues of performance of the planned activities and consequently is more relevant at the tactical level, the second question attempts to identify whether the planned activities are effective in bringing the desired progress and is relevant at the operational and strategic levels. In addition, the second question is more difficult to answer, since it by necessity involves the effects achieved by a synergy of multiple actors. This paper presents one possible approach to the assessment of mission effectiveness and progress towards the strategic objectives, based on the author's experience with supporting the assessment of the Canadian mission in Afghanistan.

The paper addresses three key issues:

> ➢ Topology and connectivity of the assessment,
> ➢ Non-linearity and emergent properties, and
> ➢ Subjective vs. objective assessment.

The first point considers the ties and relationships that exist between different levels of assessment and different stakeholders.

The second bullet outlines the fact that systems, which did not have enough time to establish equilibrium cannot be characterized by "average" quantities. For example consider the following equilibrium situation: the conditions in a closed room can be partly characterized by the average temperature (i.e. average over temperature measurements taken at various points in a room). However, if there was not enough time (or it was impossible) to establish equilibrium, the concept of average would lose its meaning. For instance, while the average of the temperature measurements between a sauna and adjacent showers can be calculated, the value does not have any meaning. It characterizes neither the conditions in the sauna nor the conditions in the showers. Similarly, it is possible to calculate average age of a family, but it does not provide any meaningful information about the ages of individual members.

The last point looks at assessment scales, selection of measures, and finding a balance between subjective and objective assessment. This paper does not attempt to fully address this issue, since

the selection of measures, and in particular human factors, present an entirely different challenge and their adequate treatment goes beyond the intended scope of this paper.

The paper is organized as follows. First the issue of self-similarity and topology of the assessment is discussed. A possible approach for a non-self-similar environment is outlined. Then the example of an "effects dashboard" developed for the assessment of the mission in Afghanistan is presented [3]. It includes the methodology for the roll-up of the assessments as an example of combining lower-level assessment to obtain a summary assessment at a higher level without averaging. Lastly, the connection between tasks and effects (and consequently measures of performance and measures of effectiveness) is briefly touched upon in the context of the topology of the assessment.

Topology of the Assessment – From Tactical to Strategic

The military traditionally structures itself into multiple organizational levels that are often loosely based on geographical (spatial) structures, in particular in a conflict environment. The highest level is strategic, concerned with the interaction with the political sphere and the larger geo-political picture. Geographically, in terms of deployed operations, the strategic level would deal with the entire theatre of operations. This is followed by the operational level that is generally concerned with a particular conflict or mission, providing direction to the deployed troops. Again, geographically, the operational level would be dealing with the appropriate area of operations. Finally, a tactical level is considered within particular operations, battles and engagements. The military structures tend, to a large degree, to be self-similar (i.e. the structure of headquarters at a higher level may be almost identical to that of its subordinate headquarters). On the other hand, unlike in Canada where the political structures are almost self-similar (elected federal and provincial parliament, municipal councils), political structures in many areas are not self-similar, and can differ significantly between different geographical levels (i.e. the structure and selection system of a village council is different from that of a national government). In some cases, like Afghanistan, there might be multiple, almost parallel, political (governance) power structures (e.g. tribal and district leaders). When such structures need to be related, it presents an additional challenge to the assessment.

A commonly used means of dealing with this challenge in assessment is to use multilayered assessment matrices, with a different layer for each organizational level. These enable the lower levels to see the big picture behind the desired objectives and effects, and at the same time they enable the higher levels of the chain of command to identify the problem areas, and perhaps explain why certain areas do not progress as expected. The use of multilayered structures has an additional advantage. Since the higher levels are usually too aggregate to provide enough details for an enemy to use, different layers of the assessment structure can potentially have different classification levels assigned to them, thus enabling seamless reporting to the public without revealing tactical details that the enemy could use to advantage.

There are two possible situations that require distinct approaches. In a situation when the assessment structure is intended to serve for multiple levels assessing the same geopolitical level (e.g. provincial, national), the structure can be self-similar, with the higher-level assessment based solely on the assessments of lower levels (Figure 1). An example of such a structure is the

effects dashboard developed for the assessment of Kandahar Province, which is discussed in detail the next section.
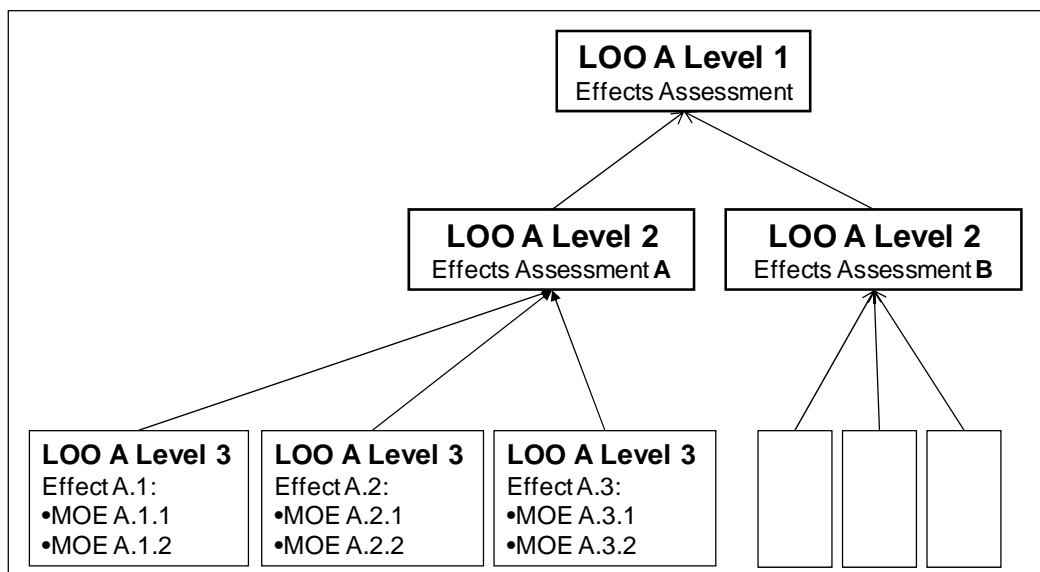

Figure 1. Schematics of the multi-level assessment in a self-similar environment.

However, when the assessment structure is intended to assess distinct geopolitical levels (from district to provincial to regional…), the self-similar structure does not apply, and the assessment at higher levels needs to consider both the lower levels, and the level-specific information (Figure 2). A possible approach to capture the behaviour of such non-self-similar systems would consist of a three-step assessment at each level as follows.

The assessment starts with the lowest level. This could be a district or a village or another structure representing the smallest entity the assessment is interested in. For further example, it is assumed that the lowest structure is the village. Different villages can use different measures/indicators, as long as these are connected to common operational effects. Based on the assessment of the villages, the assessment of district (or other higher level) is made using some pre-defined roll up. However, this is not sufficient, since it does not capture any characteristics not contained in the village assessments but relevant at the higher level. Therefore, the second step is undertaken, in which the district is assessed using the metrics relevant for that particular level and independent of the lower-level assessment. Then, in the third step, these two pictures are combined to create the complete picture at that particular higher level. This approach can be extended all the way up to the strategic assessment of the entire region. In practice this means that in the case of an assessment in such a non-self-similar environment the dashboard-like structure would need to be extended to include the room for these intermediate assessment steps.

For instance, to arrive at the assessment of the second-level effects, at first the lowest-level effects would be combined in a way that might have been similar to the one described in the following section discussing the effects dashboard (Figure 3). Then the independent assessment of the second-level effects using appropriate indicators and measures would be performed, and finally these two would be rolled together to obtain the final situational picture for the second-

level. Since the author was not directly involved in the assessment of a non-self-similar structure, no example of a real tool was available.
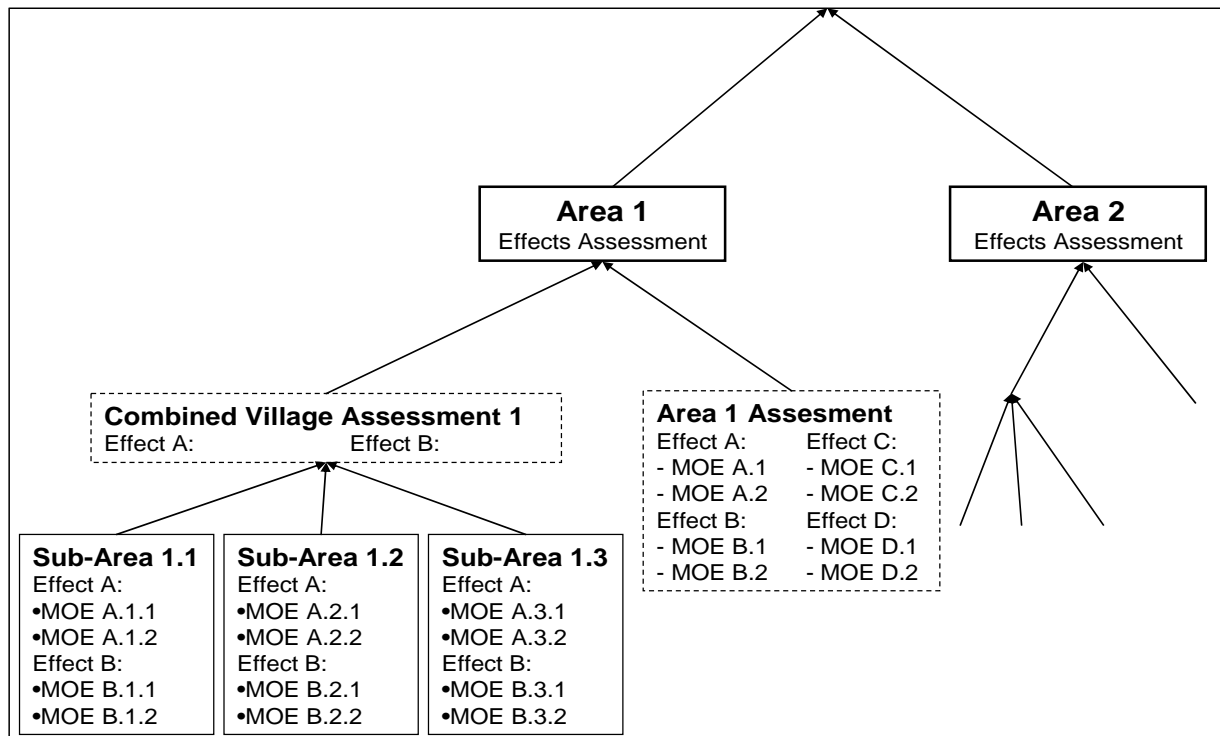

Figure 2. Schematics of the multi-level assessment in a non-self-similar environment.

Example of Self-Similar Assessment: The Canadian Effects Dashboard

This section presents an example of a self-similar assessment structure developed in support of the assessment of the Canadian mission in Afghanistan. The mission in Afghanistan is an excellent example of a complex environment. It involves joint work of multiple government departments and other stakeholders with different objectives and different organizational structures, each influencing all the others. To enable work in such diverse environment, the mission was divided into a number of lines of operation (LOOs). This can be a very useful construct enabling one to capture the distinctions in the purpose and (to a degree) in timescales. However, the LOOs are deeply intertwined, with a number of influence links. Additionally, all of the actors can influence all LOOs, whether directly or indirectly. Nevertheless, this construct can help to ensure that no area of the mission gets ignored or suppressed in the assessment. However, the LOOs are still subject to the same limitations in terms of spatial and organizational structures as the overall mission (as outlined in the previous section).

To address the reporting requirements of Canadian Expeditionary Force Command (CEFCOM), a comprehensive assessment methodology for Kandahar Province was proposed. The key characteristic of this methodology was a multilayered tree structure enabling the capture of different organizational scales, as well as multiple stakeholders. To line up with the traditional requirements, the tree was divided into branches along the three LOOs: Security, Governance,

© 2010, Small Wars Foundation

and Development. The important point was that the effects were distributed uniformly across the LOOs, and that the number of considered effects was fixed.

The key reasons behind the use of the tree structure in the presented form were as follows:

- ➢ The use of multiple layers enabled assessment relevant for different organizational structures (such as deployed headquarters, CEFCOM, Strategic Joint Staff);
- ➢ The self-similar structure enabled smooth roll up of the assessment as will be discussed below; and
- ➢ The balance in the effects across the three LOOs was intended to provide a coherent picture across all the LOOs.

After numerous discussions with stakeholders, the considered effects were given an "importance score" based on how important they were considered for the overall progress (i.e. vital, important, necessary, and desired). This enabled a more robust roll-up process, and it also provided a more valuable output highlighting the most important areas that required additional resources.

The balanced tree structure was supported by pre-defined effects scales that reduced the subjectivity of the assessment. The scales were based on the set of indicators assigned to each measure at the lowest level. These indicators then formed the basis for assigning grades during the assessment as is shown in Table 1. This approach reduced the subjectivity of the assessment. An important point needs to be made here. While the effects were assigned numerical grades, these were treated as symbols (akin to grades A, B, etc.), not as numbers. So no numerical operations (e.g. averaging) were to be performed with the grade values.

5	**Fully achieved:** The effect can be considered achieved; all of the indicators are present to a significant degree, only some minor improvements can still be made.
4	**Acceptable standard:** The effect is almost fully achieved; though most of the indicators are present to a significant degree, some improvement can still be made.
3	**Below acceptable standard:** The effect has not been fully achieved, but a noticeable improvement is observed. Most of the indicators are present, albeit to a minor degree, though improvement is still required in this area.
2	**Minor:** The effect has not been achieved; however, a few indicators are present to a minor degree, but significant improvement is still required in this area.
1	**Negligible:** The effect has not been achieved; almost none of the indicators are present and significant improvement is required in this area.

Table 1. Evaluation scale used to assign scores to the lowest level of measures.

To assess higher levels on the basis of the lower-level assessment, an automatic roll up was introduced in the structure. It was based on the assumption that since the assessment related to only one geopolitical structure at all levels, the higher levels could be assessed using only the information from the lower levels (reducing granularity while preserving structure). The algorithm accounted for the fact that some of the measured characteristics were more important

than the others, and that simple "averages" of the grades would be meaningless. The algorithm is shown in Figure 3. It is condition-based, accounting for the relative importance as well as the performance of the lower-order effects. The importance of the effects was decided by t he assessment team after consultations with the stakeholders. The roll-up process assumes that there is always at least one vital effect present in a branch. This roll-up algorithm demonstrates that it is possible to combine lower-level assessment to obtain a higher-level picture without resorting to the averages. Similar schemes might be developed for any assessment as required.
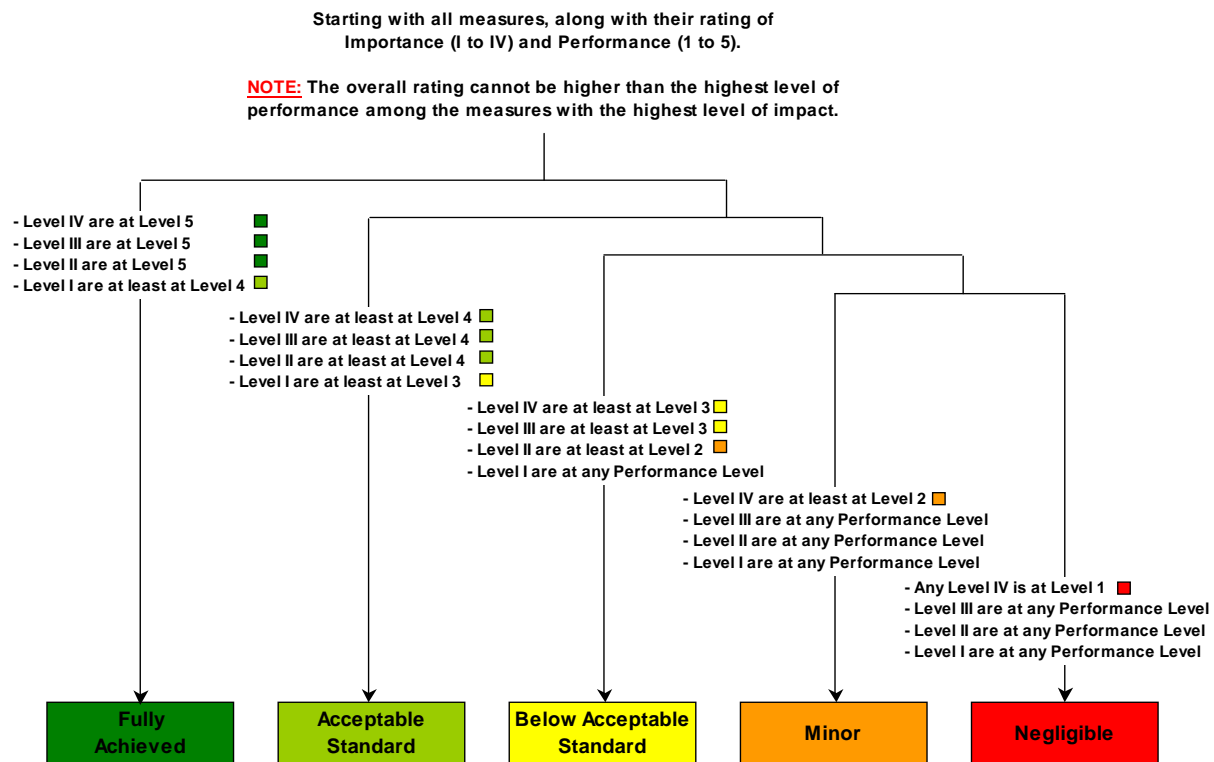
Starting with all measures, along with their rating of Importance (I to IV) and Performance (1 to 5).

NOTE: The overall rating cannot be higher than the highest level of performance among the measures with the highest level of impact.

- Level IV are at Level 5
- Level III are at Level 5
- Level II are at Level 5
- Level I are at least at Level 4

- Level IV are at least at Level 4
- Level III are at least at Level 4
- Level II are at least at Level 4
- Level I are at least at Level 3

- Level IV are at least at Level 3
- Level III are at least at Level 3
- Level II are at least at Level 2
- Level I are at any Performance Level

- Level IV are at least at Level 2
- Level III are at any Performance Level
- Level II are at any Performance Level
- Level I are at any Performance Level

- Any Level IV is at Level 1
- Level III are at any Performance Level
- Level II are at any Performance Level
- Level I are at any Performance Level

| Fully Achieved | Acceptable Standard | Below Acceptable Standard | Minor | Negligible |

Figure 2. The rollup algorithm for the assessment of higher level measures. The relative importance goes from I (least important) to IV (critical).

Tied with the roll up was the selection of "problematic" areas. These were characteristics that were deemed critical or important and did not perform too well. There were three levels of importance: those requiring immediate attention (vital areas given the score of 1), those that might require attention (important areas given 1 or critical given score 2), and those who should be attended to should there be an opportunity (important areas given score 2). These would be marked with flags of different color in order to highlight them and thus provide quick information to the stakeholders.

Lastly, the assessment of the mid-level characteristics, appropriate for CEFCOM, was conducted. The variations between assessment periods and the relationships between them could easily be assessed manually by analysts (it involved only 32 characteristics).

Performance Measurement in the Effects Assessment: Connecting Tasks to Effects

One of the challenges of assessment is that there is a natural inclination to "measure" things that are inherently easy to measure (such as a number of schools built or a number of security patrols conducted). However, the problem with these measures is that they assess how well the tasks are conducted and not necessarily whether they contribute to the desired effects.

Yet, measuring performance, as opposed to effectiveness discussed in the previous sections, can contribute to the overall assessment picture, under two conditions: first, the tasks are well defined and connected with the desired effects in the planning process (in other words, there is a reason to expect that a good performance of the tasks will lead to the desired effects), and second, measuring performance is used only in a supplementary role and is combined with other indicators to obtain the assessment of the effects achieved by the performance of the tasks. The second condition is especially important since it provides the necessary feedback for the planning process to verify that the assumptions and expectations materialize [2].

If these two conditions are satisfied, measuring performance can provide an enriched tactical picture and their connection with the effects can ensure that the tactical planning is in line with the operational and strategic objectives. If these two conditions are not met, the use of performance measurement becomes irresponsible, because it will almost inevitably lead to a skewed operational picture, and it will encourage conduct of particular tasks without considering their impact on the overall operational situation.

Conclusions

Mission assessment in a complex environment such as current conflicts in Afghanistan and Iraq present a significant challenge to the assessment community. However, a robust and balanced assessment is a necessity for the conduct of the COIN and state-building activities.

The two key points that were observed in Afghanistan are the requirement of developing an appropriate fidelity of the assessment at each level, and developing a robust methodology for assessment across these levels. Due to differences in security, governance, and economic structures across levels (e.g. from village to provincial), it might be impossible to simply roll up the lower-level assessment into a higher-level assessment. This is especially challenging for the reporting from taskforces to regional commands and higher.

It makes also challenging the assessment at lower levels, and often multiple schemes might be required to provide the necessary robustness and objectivity. At the lowest level (village or community) human factors, especially human behaviours are likely to provide the best picture. At higher levels more abstract set of indicators might be necessary to enable assessment across the communities.

Dr. Peter Dobias received his MSc degree in theoretical physics from the Comenius University, Bratislava, Slovakia and his PhD in physics from the University of Alberta in Edmonton, Alberta, Canada. He works for Defence Research and Development Canada Centre for Operational Research and Analysis. His focus is complex systems, combat modeling and

© 2010, Small Wars Foundation

wargaming, as well as effects-based operations and assessment. Recently he returned from Afghanistan where he worked as an Operational Analyst in Support of Task Force Kandahar, and supported the planning and assessment efforts of J5 Division.

References

1. Bowman, C.W., " Operational Assessment – The Achilles Heel of Effect-Based Operations", Naval War College, Newport RI, 2002
2. Dobias, P., Q. Innis, J. Caruana, Effects Based Approach to Operations in Task Force Kandahar: Incorporating Effects in Planning and Assessment Cycles, DRDC CORA TM 2009-049, November 2009.
3. Kalantzis E., P. Dobias, and D. Connell, "Assessment of Effect Based Operations Based on a 3D Approach", In Proceedings from Cornwallis XII: Analysis for Multiagency Support, Pearson Peacekeeping Centre, Cornwallis Park, Nova Scotia, 2-5 April 2007

© 2010, Small Wars Foundation