Evaluating performance of military deployment in crisis response operations

Erik de Waard and Robert Beeres

NL-ARMS 2007, xx-xx

Abstract

This paper contributes to measuring effectiveness and efficiency of crisis response operations. By asking whether it is useful to measure the effectiveness and efficiency of crisis response operations, we directly address the senior leadership of the Netherlands armed forces. We conclude, first, that operational commanders have a need for relevant indicators to analyze the effect of their actions. Second, we show that within crisis response operations it is usually the reconstruction tasks that lend themselves best for performance measurement. Other military tasks cannot, or only with great difficulty, be measured objectively. For these tasks we propose a greater emphasis on steering on input. In this way the organization gets more insight into the employable potential of people and means it can draw from for crisis response operations. Finally, we suggest that the output of the Dutch armed forces can be objectified further by relating it to the output of similar armed forces.

Introduction

Within the government, and the military domain specifically, the use of notions, such as “performance”, “effectiveness”, and “efficiency” have not been in use for a very long time. Thinking in terms of economic concepts, such as revenues, costs and profit was something for commercial enterprises, not for governmental organizations aspiring for “higher objectives” (see Drucker, 1998: 89). Moreover, for the armed forces it could be said that human lives, and possibly the continued existence of the entire state, were at stake. At such times, money should be left out of the equation! (see de Bakker, 2005: 183). Since the 1970s, with the advent of New Public Management (NPM), the use of such terms has become increasingly common practice within the government. At the moment – even within the armed forces – effectiveness and efficiency, performance-orientation, may be felt to be normal, if not something worth striving for.

The NPM movement aims at ‘lessening or removing differences between the public and the private sector and shifting the emphasis from process accountability towards a
greater element of accountability in terms of results’ (Aardema, 2005: 8; also see Hood, 1995: 94). Ter Bogt (2006:1) regards NPM as a ‘functionalist approach, in that one of the changes it proposes is to increase economic efficiency and effectiveness in public sector organizations’. Aardema (2005: 9) indicates that in the Netherlands ‘in the implementation of NPM much attention has been given to improving planning and control. Within the government NPM was realized through a project called From Policy budgeting to Policy Accountability (VBTB). In this project there are three (the “three Ws”) central questions related to planning: “What do we want to achieve? What are we going to do to achieve it? What amount of money can we spend? Related to control, there are the three so-called “H-questions”: “Have we reached what we aimed for? Have we done what must be done? Has it costs as much money as we thought it would?” In order to answer these questions a link must be made between the policy objectives (expressed in effects to be achieved), the efforts that ensue from them, and the resources to be employed.

Against this VBTB background the Court of Audit assessed in 2002 and 2003 to what extent the 2002 and 2003 budgets of all the Ministries and the 2002 annual account provide information about policy and management effectiveness. One of Court’s findings was that, ‘information about performance effectiveness in the budgets and annual accounts was absent’ (Tweede Kamer, 2003-2004: 12).

Against the same background the Directorate of Policy Evaluation (DGFC, 2006) evaluated specifically for the Defense organization the policy article Execution of crisis response operations. It looked into the actual measuring and measurability of effectiveness and efficiency of actual military deployment in crisis response operations. DGFC (2006: 20) found that on the basis of this investigation no well-founded judgments could be made on effectiveness and efficiency of the Dutch contribution to those operations. What can be concluded from this is that on a political-managerial level it is a problem to give a meaningful substance to measuring effectiveness and efficiency of crisis response operations.

The present article attempts to make a contribution to the literature on measuring effectiveness and efficiency of crisis response operations. By asking whether it is useful to measure the effectiveness and efficiency of crisis response operations, we are directly addressing the senior leadership of the Netherlands armed forces.

To this end this contribution is set up as follows. Section 2 presents a concise survey of how concepts such as effectiveness and efficiency can be related to crisis response situations. Section 3 gives a description of the set-up of the empirical study, the findings of which are presented in section 4. Broadly speaking, there are a number of senior managers who do not find it useful to measure the effectiveness and efficiency of crisis response operations, as they think the context in which they take place is too complex. In these circumstances there are so many factors of influence that cannot be controlled
by the commanders and the organization itself. The only thing that can be done to make performance transparent can be described as a sort of broadly shared self-evaluation of the mission. Other senior managers believe that measuring effectiveness and efficiency is useful and should be done better. They want criteria for assessing how they are doing their job. They feel that the Netherlands armed forces, as a reconstruction organization, should use certain criteria to map out their effectiveness. Finally, section 5 contains a summary, conclusion and discussion.

Performance: measuring effectiveness and efficiency of crisis response operations

The control of effective and efficient activity is shaped differently in government organizations than in private companies. For a business company the sale of its products or services leads to financial revenues through the price mechanism. As such, it is an expression of the effectiveness of company. In order to sell the products and services it will have to incur costs, which is an expression of the efficiency. This means that the company can steer on effectiveness and efficiency on the basis of a performance criterion: profit.

Government organizations do not seek profit and the relation between effectiveness and efficiency is often harder to identify. This is particularly so because the products or services they supply often cannot very well be defined in terms of financial revenues; after all, more often than not there is no market for the collective good they have on offer. That is why government organizations are financed through the budget mechanism, in which, after the levying of taxes, the political decision making process leads to the establishment of a budget. The organization has to use this budget as effectively and efficiently as possible to deliver goods and services. The result of this situation is that both effectiveness and efficiency - the performance - need to be measured in a different way than in private companies.

Measuring performance is aimed at an organization being directed both at “doing the right thing” (effectiveness) and “doing things right” (efficiency of efforts) (Espejo et al, 1996). Both concepts can be given different emphases in public, non-profit and for-profit organizations. Public organizations traditionally are all about effectiveness - the fire brigade, police or ambulance has to be on the spot as soon as possible. In fact, efficiency (costs) is felt to be of lesser importance in government organizations. This, however, has been changing over the past few years now, with an increased attention for costs (see Skaerback and Thisted, 2004). Berman (2006: 7) states that efficiency as a target is pursued more by profit-organizations, because success there is measured in terms of profit. It must be said, though, in this context that, without objectives, it is impossible to
strive for efficiency. In fact, the concept can be seen as ‘a stronger form of effectiveness in that it presupposes effectiveness’ (Speklé, 2003: 4). Finally, Berman (2006: 9) states that, next to effectiveness and efficiency, public organizations pursue equity as an objective, ‘providing services to all citizens, regardless of their ability to pay for such services’. Depending on the specific situation of an organization, a performance measurement instrument is designed which is, for instance, more geared to effectiveness, efficiency or equity (see, e.g., Burchell, Clubb, Hopwood and Hughes, 1980; Hopwood, 1987; Johnson and Kaplan, 1987; Kaplan and Atkinson, 1989).

Traditionally, for the Defense organization effectiveness, measured by readiness, has held pride of place, with an implicit or explicit prominence of the defense of the territorial integrity of the state, in particular the deterrence a standing army can bring to bear. During the Cold War this readiness was virtually the only criterion on which to assess performance of the armed forces (Hazeu, 1980). On the basis of the then doctrines for the Dutch defense tasks in NATO context - the defense of the North German plain against a concrete threat from the East – precise criteria could be formulated for the required military “readiness”. The norm for the intended level of readiness could be broken down in terms of operational readiness – nature and size of the required “capabili- ties” and activation terms of anything between “rapid reaction forces” and mobilizable units. The operational readiness could further be specified in proficiency requirements (according to level of training) and the personnel and material readiness on the basis of filling and “fitness for use”. Such norms and requirements, on their turn, were the foundation for the decisions with regard to numbers of conscripts to be drafted and investments in weapon systems and other materiel.

Now that the armed forces have actually been deployed in crisis response operations to an ever-increasing extent, the use of the Defense organization can no longer be primarily described in terms of operational readiness. Actual deployment requires evaluation criteria in terms of actually attained results. They should form the basis for deducing effectiveness and efficiency of the deployment. Apart from that, the requirements for general operational readiness have become more diffuse. Combat power is no longer the leading perspective in all cases. Crisis response operations can often entail a broad spectrum of police and humanitarian tasks, demanding completely different skills. There can be considerable discrepancies between the “capabilities” of the military units deployed and the appeal that is actually made on them. Present peace operations clearly illustrate the increasing blurring of military and police tasks. Besides, the new threats from transnational terrorism make it sometimes all but impossible to delineate the tasks of the Defense organization and the Ministry of the Interior.

Measuring performance - effectiveness and efficiency – of the current expeditionary armed forces can be summarized as in figure 1 (Cornielje, 2007: 16). Underlying this
scheme is the customary analysis of a production process in terms of input (resources), throughput (activities), output (performance) and outcome (effects) (see, e.g., Berman, 2006; Mol and Beeres, 2005 and Tweede Kamer, 2003-2004). In order to analyze this process, it is important to develop meaningful and measurable indicators (see for relevant input (step 1 in figure 1) and output (step 3 in figure 1): Van den Doel, 2004: 21-34).

Figure 1 shows that the “production process” of the armed forces can be divided into two sub-processes. The first of these, the steps (1) to (4) in figure 1, is directed at preparing units for readiness, the second, the steps (5) to (8) in figure 1 is directed at the actual deployment of these units. In the latter sub-process deployed units can be seen as input for the mission, whereby the eventual consequences for the country (countries) in which the mission takes place (for instance, peace) and the consequences for the country that sends the units (for instance, increased political support) form the outcome.

Figure 1 shows eight steps from input to eventual result. For the corresponding indicators for the first five steps it is (reasonably) feasible to find measurable aspects. It is supposed that with its presence in the mission area the Defense organization generates effective security as a product. The last three steps (6. throughput, 7. output, 8. outcome) relate to the actual execution of crisis response operations, the consequences for the countries in which they take place and the consequences for the countries that carry them out.

Figure 1 Schematic representation of the Defense ‘production process’
A recent survey by the Directorate of Policy Evaluation (DGFC, 2006) shows that no founded conclusions can be drawn for the second sub-process of figure 1 with regard to the effectiveness and efficiency of the Dutch contribution to those operations. The survey, directed at the policy article Execution of Crisis Response Operations, reveals that the objectives of crisis response operations have been formulated so broadly and abstractly that no judgment can be made about the effectiveness and efficiency of the policy. Nor is it possible to steer the Dutch participation in international crisis operations on the basis of those objectives.

Furthermore, the survey evaluates for ten crisis response operations in the period 2000-2004 whether the policy was executed effectively and efficiently. On the basis of the available information the Directorate indicates that for the overwhelming majority of the surveyed crisis response operations no separate SMART objectives were formulated and that it is not possible to determine precisely what the Dutch share was in attaining the mission objectives. The above means that no conclusions can be drawn as to the effectiveness of the policy.

It was also concluded that to an extent this problem cannot be solved (DGFC, 2006: 21). In the budget phase the objectives cannot be worked out “SMART-ly” because of continuous changes in the international security environment. ‘On the one hand, this can be related to the unpredictable character of the “supply” of crisis response operations as well as their course. On the other hand, the absence of “SMART-ness” is related to the complexity of the political field of force in which the decision making takes place’ (DGFC, 2006: 5). It is even indicated that on the political level any concreteness about the SMART-ness of the objectives of crisis response operations is ‘sometimes undesirable, because the responsible politicians experience it as a restriction to policy margins’ (DGFC, 2006: 6).

It is further indicated that the Netherlands in fact never acts unilaterally and always makes a contribution limited in time and size. This makes it difficult to relate one on one the Dutch contribution and the results attained a crisis response operation as a whole. This, however, gives the Netherlands some freedom of choice with regard to the selection of activities employed by the Dutch units. ‘It creates leeway for policy choices and steering of the Dutch contribution. From the usually many tasks to be fulfilled within an international crisis response operation the Netherlands can choose the ones that best contribute to the realization of its objectives and priorities’ (DGFC, 2006: 6).

DGFC (2006: 7) concludes that at the level of the policy article it is very difficult to formulate SMART objectives. At the level of the individual crisis response operations it does see possibilities for a more policy-driven approach, particularly through implementation of the VBTB concept. This will allow a more direct relation between (political) policy objectives, which are specified per operation, and the activities employed to real-
ize those objectives (DGFC, 2006: 8, 11). Below, it will be shown to what extent DGFC’s findings and recommendations (2006) are endorsed by the senior leadership of the Netherlands armed forces.

Methodology

Within the framework of a larger research into the organizational change process that the Netherlands forces embarked on after the end of the Cold War, the operational performance of the organization was studied in a broader perspective than the DGFC survey. In view of the central problem of this article, some preliminary results of this – much more comprehensive study – will be presented. To be precise, the article will zoom in on the dependent variable performance, using relevant empirical data that have been collected so far. The analysis of this variable is based on the following two central questions:

1. In how far is the Dutch crisis response effort successful?
2. In how far is the Dutch crisis response effort measurable?

In order to answer these questions two different instruments for data collection were used. First, a number of interviews were conducted with 17 representatives from the leadership of the Defense organization in order to form a well-founded idea of the senior leadership’s views on performance and measurability of the Dutch crisis response effort. Secondly, a large-scale survey was held among 3,500 officers, varying in rank between major and colonel, from the four Services. The survey concentrated on the higher officer echelons as the comprehensive research requires respondents who not only have experience and knowledge of missions abroad, but, quite emphatically, also have an insight into all sorts of organizational aspects of the armed forces. In total 1,533 persons filled out and returned the questionnaire, resulting in a response rate of 43.8%. The sample was tested for representativeness by considering the distribution of the respondents over Service and rank. The questionnaire used a validated scale for performance (Volberda, 1998). The survey confirms the reliability of this scale with a Cronbach’s Alpha of 0.77. A Cronbach’s Alpha value between 0.7 and 0.8 means that the scale used is reliable in measuring the investigated variable, in this case performance.
Results

This section consists of two parts. The first part deals specifically with the first research question, to which end the results from the interview and the findings from the survey are presented in their mutual relationship. The second part focuses on the second question on performance measuring and only uses the results from the interviews.

The operational effectiveness of the Netherlands armed forces

The senior leadership gives a rather univocal answer to the question on the success of the Dutch crisis response effort, as none of the respondents calls it unsuccessful (see table 1). In general, there is a positive assessment. Furthermore, it is striking that 5 respondents do not give an unequivocal yes or no, but leave their opinion in the middle. In fact, these neutral-voters do not deviate far from the yes-voters, as all respondents indicate that assessing the success of a crisis response operation is difficult. The multitude of factors of influence makes it impossible to make objective statements on the effectiveness of this kind of operations. Nevertheless, for 12 respondents the balance is positive, while the 5 mentioned above do not come to a definite yes or no. The arguments given to assess the success of a crisis response operation can be reduced to the five categories that will be discussed below.

<table>
<thead>
<tr>
<th>Respondent</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 Do you think the Dutch effort in crisis response operations successful?

In the first place success is linked to the fact that the Netherlands has participated in a large number of missions over the past few years. Looking back on these operations, 9 of the 17 respondents indicate that, in general, the Netherlands has realized its tasks and ambitions well. Thus, one respondent states, ‘Looking at Bosnia, Iraq and Afghanistan it can be concluded that, given our possibilities, we have been successful’. This sort of general statement is complemented with remarks such as, ‘We reach the objectives that have been set, in spite of the fact that those objectives are not always clearly demarcated, measurable and are highly political’, and ‘We see the results of what we have set out to do in a mission; our being there made a difference’.

In the second place, in the eyes of nine respondents international recognition is an expression of the success of the Dutch effort. Often heard statements in this respect are, ‘we are a valued partner in the eyes of other countries’, or, ‘the Netherlands armed forces
are taken very seriously, indeed, internationally’, or, ‘we are in the major league’, and
‘people abroad are always full of praise about the way the Dutch operate’. Nevertheless,
one respondent stresses that it is not wise to rely too much on the opinions of other
countries, as they may be expressions of (political) courtesy.

Thirdly, personal experience forms a measure for the success of the operations. All
respondents have mission experience in leadership positions and three indicated that in
their role as military commanders the success is also measured against their own stand-
ards. Thus, one respondent states, ‘I relate “success” more to the concrete improvements
I have seen during my stint in the mission area, such as the restoration of social life in
the streets and the number of new companies that were started up’. Another respond-
ent says in more general terms that he judges his operational activities as commander a
success, because with his people he has made a contribution to the creation of stability
and reconstruction in a mission area, but also because he and his people have returned
home safely.

Quality of personnel and materiel of the Netherlands armed forces is a fourth argu-
ment brought forward. Five respondents say things like, ‘The Netherlands armed forces
are professional and well-equipped’, ‘We have good, professional, level-headed people
who reach the result they have to reach’, ‘Missions have led to an enormous improve-
ment of our own armed forces’, and ‘The organizational culture as well as the mindset
of the personnel has improved, missions abroad are very much part and parcel now of
what we do’.

Political-social considerations form the fifth category. In this respect two respondents
stress the political appreciation the armed forces receive when they state, ‘In general the
Second Chamber praises the achievement of the armed forces’, and ‘Our political leader-
ship is very satisfied about our effort’. Another respondent chooses an angle with relates
to the spending of public money. He explains that with the same budget as fifteen years
ago the armed forces now realize an investment quota of 20%. Yet another respondent
points at the improvement of the image of the armed forces among the Dutch public.
According to him there is still a lot to be gained in propagating the armed forces to the
Dutch citizen, ‘Most citizens have no idea what the armed forces do’.

In answering the first question the senior leadership goes through a sort of self-evalu-
ation, in which a summing up of the various arguments leads to a preponderantly pos-
tive self-image about the operational achievements of the armed forces. Volberda’s scale
(1998), used in the survey, is also based on subjective ratings. The concrete questions
that compose this scale are given below. What is striking is that, broadly speaking, three
out of the five arguments brought forward by the senior leadership are covered by these
survey questions. A general rating about the Dutch crisis response effort is covered by
question 1; questions 2 and 3 are specifically directed at international comparison with
other, similar armed forces, and personal experiences and/or expectations come back in
question 4.
1. The Netherlands armed forces do well in crisis response operations;
2. In comparison with other similar armed forces the Netherlands armed forces do well
in crisis response operations;
3. Other armed forces are positive about our operational achievements in crisis response
operations;
4. I am proud to have been sent out or to be sent out as a Dutch serviceman/woman.

The respondents answered the questions by means of a 5-point Likert scale (1 = totally
disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = totally agree). Table 2 shows the mean
score per question for the entire population, but it also gives the total score for per-
formance. With a mean score of 3.99 for performance the result corresponds with the above-
mentioned positive sentiment of the senior leadership with regard to the operational
effort of the Netherlands armed forces.

<table>
<thead>
<tr>
<th>Question</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1528</td>
<td>1</td>
<td>5</td>
<td>4.06</td>
<td>.560</td>
</tr>
<tr>
<td>2</td>
<td>1530</td>
<td>1</td>
<td>5</td>
<td>4.03</td>
<td>.628</td>
</tr>
<tr>
<td>3</td>
<td>1529</td>
<td>1</td>
<td>5</td>
<td>3.91</td>
<td>.631</td>
</tr>
<tr>
<td>4</td>
<td>1525</td>
<td>1</td>
<td>5</td>
<td>4.12</td>
<td>.788</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>1526</td>
<td>1.67</td>
<td>5.0</td>
<td>3.99</td>
<td>.511</td>
</tr>
</tbody>
</table>

Table 2 Performance score general

In tables 3, 4, and 5 the results with regard to performance have been broken down
per Service, rank and mission experience. The most important conclusion to be drawn
from this is that within these sub-groups the positive sentiment remains unchanged.
The score on the 5-point scale lies around 4, regardless of Service, rank or times on
missions. Furthermore, these analyses yield a number of interesting significant dif-
fferences (indicated by an asterisk). In a general sense these differences do not detract
from the positive self-evaluation, but they are nevertheless worth mentioning and will
be discussed below.

<table>
<thead>
<tr>
<th>To which Service do you belong?</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Royal Netherlands Navy</td>
<td>3.83*</td>
<td>319</td>
<td>.515</td>
</tr>
<tr>
<td>Royal Netherlands Air Force</td>
<td>4.06</td>
<td>371</td>
<td>.491</td>
</tr>
<tr>
<td>Royal Netherlands Army</td>
<td>4.04</td>
<td>777</td>
<td>.511</td>
</tr>
<tr>
<td>Royal Netherlands Marechaussee</td>
<td>4.05</td>
<td>54</td>
<td>.427</td>
</tr>
<tr>
<td>Total</td>
<td>3.99</td>
<td>1521</td>
<td>.511</td>
</tr>
</tbody>
</table>

Table 3 Performance related to Service
In table 3 it is striking that the respondents from the Royal Netherlands Navy scores lower for performance than the other Services. With a final performance rating of 3.83 the opinion with regard to operational effort is still very positive, but it comparison with the other Services it is significantly lower. Probably the assessment is somewhat more moderate because in the eyes of the Navy crisis response operations have not led to drastic operational changes compared to the past. The traditional sea time-shore time structure, in which an operational sea period of 6 months is followed by a shore posting of 18 months to 2 years, remains fully intact. It could even be said that the other Services, by embracing crisis response operations, have more or less adopted this phasing from the Navy. During sea periods the ships can be given all sorts of operational tasks, sometimes under the umbrella of a specific (multi-national) mission, and sometimes not. In other words, sea time is sea time, and it does not matter much for the Navy personnel whether they are employed in a regular maritime patrol, a counter-drug operation in the Caribbean, or in CFT150 in the Enduring Freedom framework. On top of that the Navy with its maritime task holds a rather independent position within the Netherlands armed forces. Of their nature, land and air operations are more directly linked, whereas Navy operations often take place independently. If there is cooperation, this often takes place in international fleet settings. Therefore, this autonomous position makes it hard for Navy personnel to give their outspoken opinion on the crisis response operations of the armed forces as a whole.

<table>
<thead>
<tr>
<th>What is your present rank?</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAj/Lt Com/MAj Marines</td>
<td>3.99</td>
<td>815</td>
<td>.579</td>
</tr>
<tr>
<td>Lt Col/Com/Lt COL Marines</td>
<td>4.00</td>
<td>563</td>
<td>.492</td>
</tr>
<tr>
<td>Col/ Capt/Col/Marine</td>
<td>4.07</td>
<td>144</td>
<td>.535</td>
</tr>
<tr>
<td>Total</td>
<td>4.00</td>
<td>1522</td>
<td>.511</td>
</tr>
</tbody>
</table>

Table 4 Performance related to rank

Table 4 shows that officers at the level of colonel score significantly higher than the other two rank categories. This is probably due to the average level at which colonels work during a mission. Usually this group works on the cross section between the political-strategic and the military operational levels, which makes it possible for them to relate the results of military activities to the complex power play in which they take place. So, in comparison to officers who are mainly focused on the operational execution of tasks, colonels may assess performance more positively.

Table 5 clearly shows that missions have become part and parcel of the Netherlands armed forces. Although it does not fall within the scope of this article, it is worth mentioning that to date 86 % of the field officer population has been sent out at least once.
At the same time the table reveals something that was also brought forward in the interviews: the unequal division of mission burden. A relatively large group (50%) of the field officers has been only once or even not at all on a mission. This low number of missions stands in stark contrast with the high mission intensity of some functional groups, notably pilots, engineers, special forces, Marines and medical specialists. For some of these groups mission time has been reduced in order to safeguard sustainment in spite of the “shortage”, for example, Apache and F-16 pilots, who are sent out for two a period of two months. However, this tailor-made approach is not possible for all of these groups, which results, for instance, in engineers being faced with a high mission burden as well as regular mission periods of four to six months.

<table>
<thead>
<tr>
<th>How often have you been on a mission?</th>
<th>Mean</th>
<th>N</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 times</td>
<td>3.92</td>
<td>207</td>
<td>0.438</td>
</tr>
<tr>
<td>1 times</td>
<td>4.04</td>
<td>556</td>
<td>0.514</td>
</tr>
<tr>
<td>2 times</td>
<td>4.01</td>
<td>416</td>
<td>0.532</td>
</tr>
<tr>
<td>3 times</td>
<td>3.97</td>
<td>206</td>
<td>0.531</td>
</tr>
<tr>
<td>4 times</td>
<td>3.96</td>
<td>84</td>
<td>0.469</td>
</tr>
<tr>
<td>5 times</td>
<td>4.00</td>
<td>57</td>
<td>0.531</td>
</tr>
<tr>
<td>Total</td>
<td>4.00</td>
<td>1526</td>
<td>0.511</td>
</tr>
</tbody>
</table>

*Table 5 Performance related to mission experience*

Furthermore, table 5 shows that respondents who have no mission experience score significantly lower on performance than those who do. This result may be explained by another aspect the senior leadership brought forward during the interviews. With the new expeditionary crisis response ambitions of the Netherlands armed forces the necessity for inter-service cooperation has increased enormously. Not a single Service is able to conduct an operation completely independently, from planning to redeployment. Although the cooperation is improving by the day according to the interviewees, there are regular conflicts of interest and differences of opinion on a managerial level during the planning and preparation of a mission between the various Services. Once the composition of the unit that is going to be sent out is complete, other forces gain the upper hand, as on the operational level the tasking of a mission will lead to one common interest. The ensuing task culture minimizes the problems with cooperation. Indeed, when an operation takes place in hostile and dangerous circumstances, cooperation is seen as one the most important conditions for survival. In short, in the preparation phase of a mission there is usually a lot of fuss, but once on a mission this is replaced by a professional culture, in which the participating units realize they need each other and mutual understanding increases. The respondents with mission experience have been in this
positive task culture first hand, which may express itself in their higher score compared to the respondents without this experience. The latter group cannot boast any practical experience and because of that will probably make a more balanced assessment founded on their knowledge of the much less harmonious planning and preparation phases of missions.

**Measuring operational performance**

From the interviews as well as the answers from the questionnaire it can be deduced that the average Dutch field and general officer assesses the performance of his own organization as very positive. Still, this result cannot negate DGFC’s assertion that no well-founded conclusions can be drawn on the effectiveness and efficiency of the Dutch contribution to crisis response operations. Both cases, interviews and questionnaires, concern self-evaluations, which lend the aspect of effectiveness a preponderantly subjective trait and cause a lack of objective criteria to base well-founded statements on efficiency. Conversely, the question was posed to a very large, representative group of military functionaries who know what they are talking about and who have actual operational experience. It would, therefore, be too facile to disqualify their assessments as too subjective and therefore useless. What is interesting is to see why objective criteria are lacking and whether it would be useful to spend more time, money and effort on making operational performance of the armed forces measurable. In order to do this, it is opportune to return to the in-depth interviews. The second central question was divided into two sub-questions, which were subsequently presented to the top of the Defense organization (see tables 6 and 7 below).

<table>
<thead>
<tr>
<th>Respondent</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Focusing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 6 Do you think it is possible to measure operational performance?*

<table>
<thead>
<tr>
<th>Respondent</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 7 Do you think measuring operational performance is desirable?*
In answering the question whether it is possible to measure operational performance, broadly speaking, four answer categories emerge. The yes-voters stress that in practice the armed forces have already been acquiring experience in measuring operations. The no-voters are of the opinion that it is precisely the complex, often political, power field, within which crisis response operations take place that makes the measuring process impossible. The third group does not give a clear answer either way, but these respondents stress the difficulty of measuring operational performance, because of the complex mix of influencing factors. The fourth group holds that measuring operational effectiveness is only partially possible and that the armed forces should focus on those specific aspects that lend themselves for measurement. In boxes 1, 2, 3, and 4 the answers categories are further explained by the actual statements of the respondents.

**Box 1: Measuring operational performance is possible and we are already doing it.**

Since the 1990s the armed forces have been laying more and more emphasis on achieving concrete objectives, such as the number of returned refugees, the reconstruction of employment and the number of confiscated weapons.

As Dutch armed forces we are not so good yet at measuring, but we should be improving. That is why probably a TNO team will go along to Uruzgan to carry out effect measurements. Effects that can also be measured are, for instance, price development of food, the number of riots, and explosions, et cetera.

We still do far too little output measuring of a mission, in spite of the fact that many KMAR tasks can be measured very well. For instance, the number of trained police personnel, measuring whether norm techniques are applied, the number of street incidents, the extent of cooperation with other police organizations (e.g. EUROPOL), the extent of corruption, and SMART quality plans.

As commander of a unit on a mission I wanted to make my performance measurable and I looked for criteria to gauge the performance of the unit by, such as reconstruction of public life, interviews with war lords, tribal chiefs.

In adopting the PRT concept we have embarked on a strategy directed at making our effectiveness measurable. We formulate “milestones” with regard to reconstruction activities and security sector reform tasks, etc. The criteria that we set up for this are even processed into a template for everyone within the ISAF operation.
### Box 2 Measuring operational performance is not possible

In my view measuring the operation is impossible. You can measure on different levels, which makes the results level-bound and subjective. Performance itself is not measurable, because when the shooting starts, different rules apply. Rules that cannot be phrased in management, such as efficiency and effectiveness. I also do not see the point of measuring from a so-called beneficiary perspective. You know beforehand that when you come to a very poor country with a bag full of money, the local population itself will not react objectively either.

In general the operational objectives of missions are quite ephemeral, vague and impossible to objectify and consequently hardly measurable. Creating stability in Bosnia is an important task of EUFOR, but has Bosnia really become more stable and how should that be measured? The same questions can be asked with regard to ISAF’s main task, the creation of a “safe and secure environment”. In contrast to this there is the question whether it really matters so much that the performance of the armed forces is so difficult to measure. Defense is a producer of non-events. Every day they produce a new day of peace and security. The Dutch population hardly notices this and that is as it should be. Perhaps we should be glad and accept that our objectives cannot be made measurable.

Making missions measurable has little purpose. The public opinion is often more important than the real operation, many aspects of which are not really perceived by outsiders. Often small errors and problems are magnified and widely discussed in the media. On top of that it is very difficult to measure an operation by objective criteria. Different players can all have different opinions.

### Box 3 Making operational performance measurable is a very difficult process

Quantifying effectiveness is very difficult. The number of confiscated weapons is not a guarantee of the unit involved.

Political preconditions limit the possibilities but are detached from the task execution of a specific military unit. The fact that insecurity is increasing in Afghanistan cannot be shifted on to the task execution of our servicemen.

Making performance measurable is difficult in case of a crisis control mission. On the one hand, it is a good thing to be able to give account of what you have done to reach your objectives. On the other, it is very hard to prove that you have employed the available resources effectively. Can you map the wellbeing of people? Can security be measured adequately?
Box 4 Making operational performance measurable should be directed at specific aspects

Measuring output? In general, I say “yes, because certain things can be made measurable very well, such as skill at arms, driving skills, level of training, etc.

What you do in an operation cannot be made fully measurable. When the number of ambushes decreases, does that mean it is safer? Conversely, it is possible afterwards to systematically analyze the preset tasks, such as disarming, demobilizing, elections, number of policemen trained, establishing a demarcation zone.

Certain things can be measured very well. For instance, the democratization process that can be measured with elections as milestones. For other things it is questionable whether they can be measured.

Making a military operation measurable in itself is not useful. The number of rounds fired or the number of ambushes prevented, does not say very much about the success of a mission. In my view, it is more about measuring reconstruction activities, based on the principle that the military operation is subordinate to the reconstruction. The number of policemen trained or the successful organization of elections says more about the effectiveness of a mission.

In spite of their equivocal reaction to the first question on performance measurement, eleven out of seventeen respondents answer positively to the follow-up question (see table 3). One of them summarizes this shared opinion, on the desirability of performance measurement in a military-operational context, very well when he states, ‘The tendency for making our performance measurable must be continued. Not only does it ensure structure in our operations, it is also a continuous reminder of our primary task’. Apart from these proponents of performance measurement, there are three respondents who feel that making operational performance measurable is not a good idea. Two of these no-voters (respondents 13 and 17) emphatically base their opinion on their answer to question 2. According to them, measuring operational performance is simply not wise, in view of the aspect of violence, which will always come into play, and the complex mix of influencing factors. The third no-voter, respondent 15, introduces an interesting new perspective with his argument with respect to performance measurement. He states,

‘I think that we must accept that military operations cannot be made measurable. My objections focus much more on the effectiveness of the entire armed forces. Crisis response operations are our core-business and yet we only send out 10 % of our personnel annually, and only 15 % of our money is spent on this main task. Sending out about 5,000 persons a year may meet our ambition level, but are we not making things a bit too easy for ourselves?’

By applying the Ter Beek norm less rigidly, by benchmarking ourselves better with other countries that have similar armed forces, in quantity and quality, by applying
more rigorously the readiness tables and by reaching a higher effectiveness from certain weapon systems, he believes, the total effectiveness to be gained from the armed forces can be increased enormously. With this argument, however, the idea of performance measurement shifts from the output to the input side. In other words, it is not so much the output of actual military operations that matters, but rather the input of a greater number of ready units.

Summary and conclusion

DGFC (2006) states that no well-founded statements can be made about the effectiveness and efficiency of crisis response operations due to the absence of SMART- objectives. Complementary to DGFC (2006), the present contribution has gone into the effectiveness and efficiency of the Dutch crisis control effort. The DGFC report has largely come into being by analyzing the evaluation accounts of ten different missions. This paper is based on concrete statements and opinions of field and general officers from the defense organization. The aspect of effectiveness was investigated by means of the question to what extent the Dutch crisis response effort was successful. Efficiency was analyzed by looking at the extent of measurability of the crisis response effort. The first question was investigated quantitatively as well as qualitatively. 17 representatives of the leadership of the armed forces were interviewed in order to get a well-founded answer to the question. By answering, they went through a kind of self-evaluation, in which the combination of a number of arguments eventually led to a preponderantly positive self-image about the operational performance of the armed forces. Apart from the interviews with senior managers of the armed forces, the operational performance was also investigated quantitatively. In a survey more than 1,500 field officers, from all Services, stated their opinion on the operational performance of the armed forces during crisis response operations. With a median score of 3.99 on a five-point scale the findings of the survey support the above-mentioned positive idea of the senior management. In general, it can be concluded that field officers as well as general officers assess the operational effort of their own organization as effective.

The second questions focused on the efficiency aspect. It was investigated qualitatively only among the group of 17 senior managers. Although this is a difficult process, the majority is convinced that making crisis response operations measurable is worth the effort. Especially in times when the armed forces are deployed more often in specific reconstruction missions, the military commanders increasingly feel the need for usable indicators to make the effect of their activities transparent. In contrast to the overall mission objectives, often couched in vague language, the underlying reconstruction tasks of
crisis response operations usually offer enough substance to be translated into concrete targets or indicators. Examples are the planned rebuilding of public governance in a crisis area, by aiming at re-establishment of the judiciary, the organization of elections and the training of police officers and security forces. Over the past few years the armed forces have made some progress in this respect, but the process will have to acquire more profundity of content in the near future. This military-operational vision with regard to performance measurement is complementary to one of the conclusions from the DGFC report, making a win-win situation a possibility. The report proposes linking the VTBT concept to the level of individual crisis response operations, the underlying reason being that at this level, too, a translation into concrete, measurable activities can be attained.

Taken together, this research, in contrast to the DGFC’s findings, justifies the conclusion that it is possible to give an assessment on the effectiveness and efficiency of the Dutch crisis response effort. The fact that more than 1,500 officers are so satisfied with the performance of the armed forces is an indication that the organization is ready for its operational task. It is also true that the choice for a self-evaluation method robs the judgment of a strong foundation. The disadvantage of this is that the aspect of effectiveness acquires a preponderantly subjective ring and that objective criteria are lacking to make concrete statements about that effectiveness. In this context, the authors therefore support the general recommendations of the DGCF report, which urge that the Dutch crisis response effort be objectified. The authors’ argument for this, however, is less based on the managerial principles of performance measurement than on its military-operational advantages.

First of all, the research shows that operational commanders have a need for relevant indicators to analyze the effect of their actions. The organization has already taken the route towards operational effect measurement, but the process is still in its infancy and needs to be developed further. It must be remarked here that, in view of the complex political power play within which crisis response operations take place, this should not be solely the responsibility of the armed forces. In this framework seeking a narrower cooperation with the Ministries of Foreign Affairs and Overseas Development is to be recommended.

The research shows further that within crisis response operations it is usually the reconstruction tasks that lend themselves best for performance measurement. Although conditional for the reconstruction tasks, the other military tasks cannot, or only with great difficulty, be measured objectively. For these tasks the authors propose a greater emphasis on steering on input, e.g. degrees of filling, readiness of weapon systems, and the level of skill at arms, driving, physical fitness and operational training of personnel. By steering more emphatically on input, the organization gets more insight into the
employable potential of people and means it can draw from for crisis response operations. If it looks beyond mere readiness of operational units, it may even be possible spread out the mission burden more evenly within the organization. Furthermore, it is important in this respect that education, exercise and training programs, maintenance programs and logistic processes are scrutinized in order to gather usable management information.

Thirdly, the authors feel, the output of the Dutch armed forces can be objectified further by relating it to the output of similar armed forces. This may be a possible task for NATO. By establishing an auditing body within the alliance, directed at scrutinizing operational readiness, the member countries would get a periodic independent assessment about their performance.

References
Hazeu, CA (1980) Prestatiebegroting op proef, een tussenstand. *Openbare Uitgaven*, 12


