

## PERFORMANCE INDICATORS IN THE DUTCH DEPARTMENT OF DEFENCE

NICO P. MOL\*

### INTRODUCTION

Since the 1980s throughout the Western world, attempts have been made to introduce more decentralized types of management control in government bureaucracies. Decentralization of competences and responsibilities within government organizations should improve efficiency in the public sector — specifically in 'core business' activities, where access to the market mechanism or transfer to independent agencies is blocked. Generally, such a decentralization is considered to involve a substitution of performance controls for the input controls traditionally applied. Organization units might become 'self-managing' with respect to the utilization of their resources, when they could be made effectively responsible for the results obtained.

In the Netherlands a reform labeled *contract management* has been introduced into a large number of government organizations to realize this objective (Mol, 1989). Thereby targets for production are based upon explicit agreements between central management and decentral organization units. By mutual consent, then, those targets would constitute reliable benchmarks to judge performance on the basis of outcomes.

Responsibility accounting should provide the techniques for performance measurement, required for the implementation of this contract management. Specifically, the development of indicator systems for the activities of organization units was expected to supply government with the instruments needed. Evaluation of efficiency in terms of input-output ratios calculated from these indicator systems should create an adequate substitute for the signals of the price mechanism elsewhere in use.

The idea of contract management has triggered the development of indicator systems in many government organizations. Rarely, however, has this development reached the stage of full implementation. Indicator systems characterized by unambiguous definitions of the indicators applied, with

\* The author is Professor of Public Financial Management at the Royal Military Academy and Associate Professor of Public Finance at the University of Twente, The Netherlands. Thanks are due to Peter Boorsma for helpful comments.

**Address for correspondence:** Nico P. Mol, Faculty of Public Administration, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands.

precise target values specified and actual values regularly monitored and reported, can as yet be considered exceptional (cf. Sorber, 1993).

In this paper one such an advanced indicator system will be examined. It will be noted that the system — notwithstanding the scrupulous measurement of a large number of indicators — did not result in performance controls which could effectively replace the input controls actually used. The case study presented may thus draw attention to some additional conditions which those indicator systems should satisfy. Specifically, we will discuss the *relevance* of the individual indicators and the *coherence* of the indicator system as a whole with respect to the objective of performance evaluation. Thereby a model for the construction of coherent systems of relevant indicators will be developed and applied to the case examined.

#### OBJECT OF THE CASE STUDY

The National Logistic Command (NLC), within the Department of Defence responsible for general logistic support and maintenance of material equipment on behalf of the Dutch land forces, will serve as the object of our inquiries. This organization is known to be well advanced in the application of responsibility accounting to its activities (as may generally be the case in this type of organization, cf. Euske and Lebas, 1994).

Since 1991 indicator systems have been measuring the performance of NLC and its subunits. These systems are contained in so-called TAPRO's ('task programs'), defining performance indicators and specifying (yearly) target values for these indicators. Periodically, actual values for the indicators are submitted in TAPRO evaluation reports (for a more extensive treatment of the ideas underlying the TAPRO system see Mol, 1995).

In the task programs, activities are clustered in so-called functional areas. Within these functional areas performance indicators are specified and target values are set, according to the format set out in Table 1.

However, within NLC the question has been raised, does the system in fact provide any guidance to management control at all? Actual feedback of

**Table 1**

| Functional | Performance Indicator |                         | Target Value | Reporting     |
|------------|-----------------------|-------------------------|--------------|---------------|
| Area       | Name                  | Definition<br>(formula) |              | Requirements* |

*Note:*

\*E.g., for some indicators actual values are reported every four weeks, for other indicators only incidentally.

TAPRO evaluations to decision making is being sensed as quite negligible. In particular, the intended contract management — to be built upon the task programs — has not yet materialized. Since the introduction of the TAPROs in 1991 virtually no reduction of input controls has been realized. Thus, we may ask which deficiencies in NLC's management system might explain this disappointing outcome.

In evaluating the performance controls applied in the TAPROs, it should first be admitted that external causes may be held partly responsible for this outcome. Notably, NLC has been subjected to rigorous reorganizations during the whole period of the TAPRO system's existence. Nevertheless, a closer look at the TAPROs reveals some weaknesses in the indicator systems themselves, which may explain their limited value as a tool of management. Specifically two deficiencies of the TAPRO system will be discussed: (1) a lack of relevance of individual indicators, because of an emphasis on secondary rather than primary processes in performance measurement; and (2) a lack of coherence of the indicator system, because relations between the indicators are insufficiently determined. We will focus upon one unit of NLC, 750 Maintenance Command (a regional unit for general maintenance), to provide more detail to the study.

#### *Relevance: Primary and Secondary Processes*

Performance indicators for organization units will refer, in principle, to the services which those units deliver. Problems of performance evaluation may sometimes necessitate measurement by proxies — based upon input or process rather than output characteristics of production — but need not deflect evaluation from the primary processes of production itself.

However, a classification of the indicators for 750 Maintenance Command in areas belonging to primary processes (maintenance activities) and secondary ones (related to internal management of the unit) reveals that in the TAPRO system emphasis is in fact being given to the last category. In TAPROs for 1993 and 1994 respectively, the numbers can be counted as in Table 2.

In the category of secondary processes we may encounter indicators like 'sick-leave' and 'vacancies' and a variety of indicators related to documents to be prepared: budget proposals to be submitted in time, plans and reports with respect to training programs or the reduction of environmental impacts of production etc., including the preparation of next year's TAPRO itself. Usually, these documents are just reported as available or not available (in time), as performance implied in their preparation cannot be measured otherwise.

Such indicators do not, however, measure performance in services delivered. They may rather be characterized as monitoring devices with

**Table 2**

|      | Total Number of Indicators | Indicators with Respect to: |                     |
|------|----------------------------|-----------------------------|---------------------|
|      |                            | Primary Processes           | Secondary Processes |
| 1993 | 60                         | 20                          | 40                  |
| 1994 | 49                         | 22                          | 27                  |

respect to internal management, having as their main purpose to reduce the need for 'ad hoc' intervention by superiors. As such the indicator systems do not comply with the objectives of contract management purported at their development.

#### *Coherence: Indicator Sets and Indicator Systems*

Decision making can rarely be guided by values reported for single indicators without any reference to other variables related to them. Deviations of actual indicator values from targeted ones as such, do not give a clue to any action to be taken in response. They should be explained first by connecting them to other variations observed, to determine the direction in which adjustments would have to be made.

In the indicator system developed for 750 MC, however, this coherence is found to be wanting. In the evaluation reports, values for the indicators are just being collected without any attempt to derive some general conclusion with respect to performance from them. Listings of projected and realized indicator values are only supplemented by detailed comments for individual indicators separately. Thus, the reports are presented in a line-item format as shown in Table 3. (Two indicators are added as examples.)

This lack of coherence of the TAPRO indicator systems may also explain the very limited use of performance controls based on the evaluation reports so meticulously prepared. Obviously, the indicators in these reports are not connected to some underlying performance criterion from which some meaning for their variations might be derived. Deviations from the reported targets are intended to provide 'signals' for decision making, but without any clue as to what they are supposed to signalize, decisions cannot be based upon them.

From these conclusions we turn to the question of how the additional conditions of 'relevance' and 'coherence' of indicator systems can be operationalized. We will try to find an answer to this question in a concept of value for money auditing to be developed in the next section.

**Table 3****Functional Area: Specific Maintenance Projects**

| Performance Indicator                   | Projected Value | Realized Value | Comment  |
|---|-----------------|----------------|--|
| Productive Labour Hours                 | 23000           | 21013          | Less because of transfer of labour to other projects |
| Percentage of Projects Executed in Time | ≥ 90%           | 94%            | None   |

*Note:*

As the reports are confidential, the figures presented are fictitious.

## VALUE-FOR-MONEY AUDITING

In for-profit organizations indicator systems for performance evaluation are derived from the ultimate objective of profitability. The classic example is provided by the DuPont Chart, in which the single profit criterion is being expanded — possibly indefinitely — into logically connected indicators for variables affecting this ultimate criterion.

We may consider then, if for this auditing concept, a counterpart applicable to non-profit organizations (or 750 MC in particular) could be established. This question has been raised specifically in value-for-money auditing, 'value-for-money' being the nonprofit alternative to the profit criterion in the for-profit sector. In the application of value-for-money auditing to the public sector a framework for indicator systems should be based ultimately upon a specification of social objectives pursued on the one hand ('value') and budgets allocated on the other ('money'). So we may define value-for-money as the ratio of social outcomes (effects) to budgets authorised for some purpose. In a formula:

$$E \text{ (effect)}/B \text{ (budget)}.$$

In value-for-money auditing as developed in the 1980s, and promoted especially in the UK, this single ultimate performance criterion is dissected into three components: economy, efficiency and effectiveness (Metcalf and Richards, 1990, pp. 28–30; and Carter, Klein and Day, 1992, pp. 35ff). However, substantial confusion remains over the precise meanings of these three concepts (in particular with respect to the concept of effectiveness, cf. e.g. Pollitt, 1986). In this paper the following definitions will be adopted:

- economy is the ratio of (physical) inputs used up to budgets spent  $I/B$

- efficiency is the ratio of outputs obtained to inputs used up  $O/I$
- effectiveness is the ratio of effects (impacts) realized to outputs obtained  $E/O$ .

The three components are logically related to each other and to the underlying value-for-money concept, because evidently

$$E/B = I/B \times O/I \times E/O,$$

(alternative definitions presented by e.g. McKinney, 1986, pp. 280–281, and Metcalfe and Richards, 1990, p. 29, fail to establish the required logical connections between the concepts involved).

To construct indicator systems according to this framework relevant indicators have to represent

- |   |   |   |
|---|---|---|
| <ul style="list-style-type: none"> <li>● budgets allocated</li> <li>● inputs used up</li> <li>● outputs obtained</li> <li>● effects realized</li> </ul> | } | with respect to the activities involved |
|---|---|---|

To constitute a coherent system, any set of such indicators should fulfil two requirements for its application in value-for-money audits:

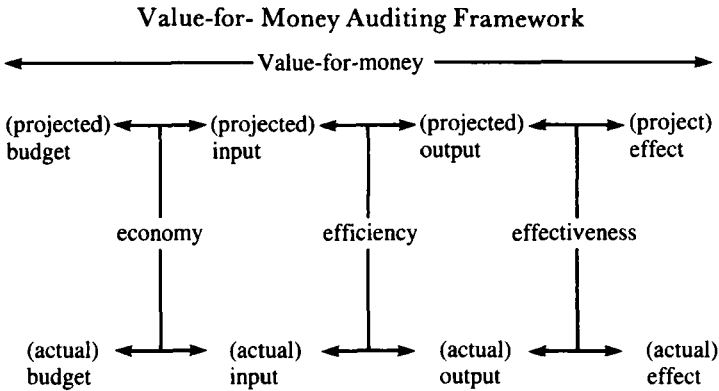
1. *Consistency*: all indicators should measure characteristics of an identical cluster of activities. E.g., an evaluation of 'economy' based upon a budget and input indicator will fail, if not all inputs used are accounted for in the budget (for instance in case of indirect costs) etc.
2. *Completeness*: performance evaluations in terms of economy, efficiency and effectiveness should contain indicators for all variables specified in the framework, as required for the calculation of the ratios by which these criteria have been defined.

Indicator systems satisfying these requirements may produce definite judgements of economy, efficiency and effectiveness and their contribution to the overall value-for-money objectives. Thus they may constitute a tool of management which effectively directs attention at deficiencies in performance and at actions to be taken in view of them.

Targets for the relevant (budget, input, output and effect) indicators are then derived from performance standards for economy, efficiency and effectiveness. They are not immediately — and from the point of view of performance evaluation arbitrarily — set for those indicators themselves. Thus value-for-money auditing may proceed by comparing actual and projected values as depicted in Figure 1.

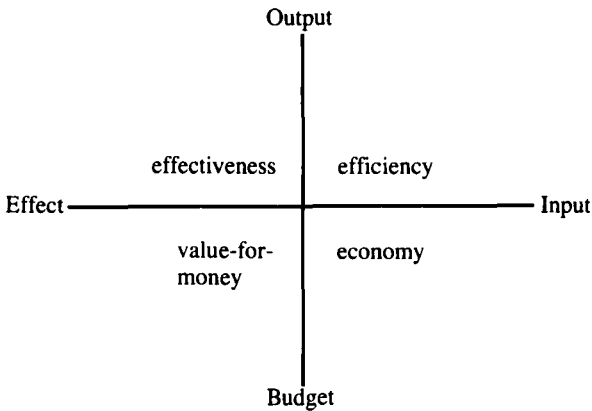
The dual evaluation involved in these comparisons (dissection in components of projected and actual values) may be facilitated by a graphical representation of the model in which the four types of indicators are measured

**Figure 1**



**Figure 2**

**Graphical Representation of Evaluation Model**

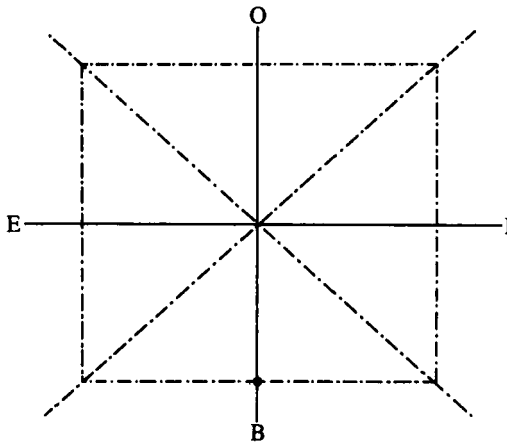


along separate axes (adapted from Bouckaert, 1990 and 1991, and Pedersen, 1977). (See Figure 2.)

Starting from some arbitrarily chosen point on the B-axis to represent the authorised budget, all other axes may be scaled according to standards set for the three performance criteria. Depicting standard economy as a 45 degree line in the fourth quadrant, the target value for inputs used up in spending that budget may be determined on the I-axis. Similarly, with standard efficiency as a 45 degree line in the first quadrant we derive a point on the O-axis for the output matching this targeted input value. Again, from standard effectiveness a corresponding point on the E-axis may be derived. Thus we have the situation depicted in Figure 3.

**Figure 3**

## Setting Standards for Performance Criteria



All ratios set for the performance criteria are implied in a connection of indicator values represented in the model by a square. Performance evaluation may then proceed by comparing any rectangle depicting actual values in the model with this square.

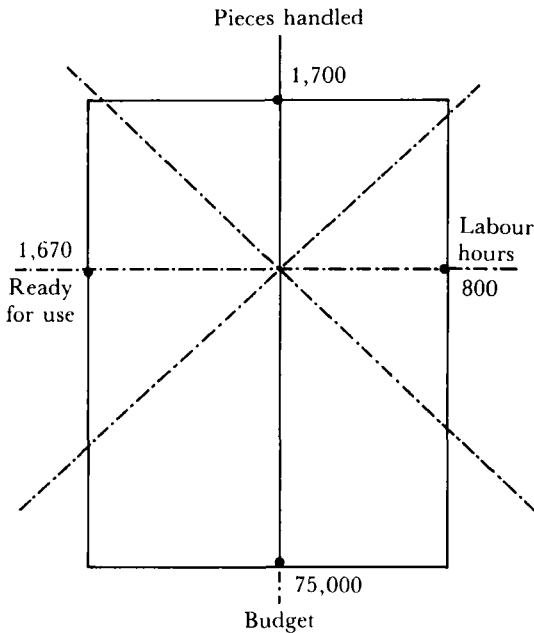
## APPLICATION TO 750 MAINTENANCE COMMAND

The application of the model to maintenance activities may be clarified by a (fictitious) numerical example. Let us suppose a budget of \$50,000 has been authorised for inspection and repair of a certain type of equipment. For budget spending three performance standards have been specified: (labour) input may cost \$50 an hour, on average two pieces of equipment should be handled in one hour and 98% of inspections and repairs should be successful. Let performance measurement then result in the following actual indicator values: budget spent \$ 75,000, (labour) inputs used up 800 hours, outputs obtained 1,700 pieces handled, effects realized 1,670 pieces ready for use. This example may then be depicted as in Figure 4.

Value-for-money may be considered insufficient in this case, to be attributed to the low I/B ratio in particular (only partially offset by favourable outcomes with respect to efficiency and effectiveness). Possibly, in this case input payments should be scrutinized, because higher prices paid for labour would not be accounted for by proportional increases in labour productivity.

Obviously, such straightforward assessments will rarely be fully justified. Deficiencies in the validity or reliability of indicator measurements usually will call for more cautious applications of the model, in which alternative



**Figure 4****Performance Evaluation of Maintenance Activities**

interpretations of measurement results will have to be taken into account (e.g., a relative decline in productive labour hours — the 'job time/shop time' ratio — may constitute an alternative explanation for the inadequate performance on the 'economy' criterion). However, the model presented may prove useful in such circumstances too, because it provides a simple visual frame of reference to ponder such alternative interpretations simultaneously.

Upon this evaluation model the performance controls required for contract management with respect to 750 MC can be based. Applying the model to the TAPROs of this organization unit, we may thus consider to what extent a coherent system of relevant indicators could be derived from these task programs to generate those controls.

An overview of available relevant indicators in the TAPROs should enable us to assess the deficiencies of 750 MC's present indicator system. In those TAPROs, for eight types of maintenance activities two indicators are calculated: labour hours spent and percentages of projects delivered in time. Budgets are not allocated to the eight types of activities separately. In 750 MC's budgets, expenditure is classified by category of resources into 'personnel' and 'material'. Budgets are thereby restricted to controllable costs, implying especially with respect to personnel expenditure that only a fraction of full costs is being covered (numbers and salaries of employees are

fixed in NLC centrally). From this overview some major incompletenesses and inconsistencies of the indicator system may be evident.

*Incompleteness* of the system: Budget information is only partially available, because costs cannot be allocated to the different types of maintenance activities on their own. As an input indicator we have labour hours. Outputs and/or effects are indicated by one variable 'percentage of projects delivered in time' only. The interpretation of this indicator is not quite clear: on the one hand the 'percentage' of projects does not give a clue to the work load involved, and on the other hand, projects 'delivered in time' do not indicate if pieces of equipment or vehicles delivered will be ready for their intended use by 750 MC's customers.

*Inconsistency* of the system: The input and output indicators, both being related to the same maintenance programs, are fairly consistent. But the input indicator is not connected to 750 MC's budget at all. Labour hours inputted refer to resources allocated to the organization unit by NLC centrally. Over 90% of the unit's own budget is spent on material expenditure, which is not reflected in this input indicator.

These incompletenesses and inconsistencies of 750 MC's present indicator system obviously impede its application in performance evaluation. Nevertheless, our assessment of these deficiencies suggests that relatively slight amendments and extensions of the indicator system may suffice to build at least some value-for-money audit upon it.

#### CONCLUSIONS

Indicator systems should fulfil the requirements of relevance and coherence to generate effective performance controls. For these requirements, specifications can be derived from value-for-money auditing. Applying a value-for-money auditing model to our case study, we have examined how these requirements may be given empirical content.

However, relevance and coherence of the indicator system actually in use for the maintenance activities of the organization unit considered here have been found wanting — notwithstanding the substantial efforts spent in the development and application of this system. We may ask then, what conclusions can be drawn from this apparent failure to realize the potential for performance control which the indicator system contains.

The traditional orientation of management control in military organizations at 'command relationships' may give us a clue in this assessment. Contrary to the idea of contract management formally endorsed in the Dutch land forces, management control thereby in fact remains focused — to some extent at least — on hierarchical subordination of commanding officers to their superiors. This orientation is reflected in the application of indicators which are irrelevant and incoherent from the point of view of performance measurement.

Targets set for commanding officers may be only partially directed at performance of their organization units as such. Alternatively, those targets may be aimed at the enforcement of some compliance to central authority or regulations, and thereby focus on the way in which the units are managed internally (specifically in times of peace). Thus, indicators related to internal management will express a factual intention to apply input rather than performance controls.

Hierarchical subordination implies, moreover, an emphasis on submission to given instructions. In such a setting target values for performance indicators may be viewed as orders rather than objectives. This conception of targets as orders to be obeyed may clarify the neglect of coherence in indicator systems. Accounting for deviations from these targets will refer to the indicators — orders — per se, not to any underlying performance criterion embedded in the indicator system. Management reports will thus focus on those deviances directly rather than on their explanation in relation to any criterion of economy, efficiency or effectiveness.

Ultimately then, relevance and coherence of the indicator systems applied enable us to judge the extent to which a purported commitment to performance control (or contract management) in organizations is being given instrumental evidence. While performance controls may be officially acclaimed, managers in many organizations — civilian as well as military — in fact continue to favour input controls instead. Deficiencies in the indicator systems used may draw attention to this fact.

#### REFERENCES

- Bouckaert, G. (1990), *Productiviteit in de overheid* (Leuven, in Dutch).
- (1991), 'Efficiency and Effectiveness, a Conceptual Niche', Paper presented at the Public Productivity Working Group of the International Institute of Public Administration (Brussels, November).
- Carter, N., R. Klein and P. Day (1992), *How Organizations Measure Success. The Use of Performance Indicators in Government* (London: Routledge).
- Euske, K.J. and M.J. Lebas (1994), 'Performance Measurement for Maintenance Depots', Submission for the European Accounting Association Annual Meeting (Venice, April).
- McKinney, J.B. (1986), *Effective Financial Management in Public and Nonprofit Agencies* (New York).
- Metcalfe, L. and S. Richards (1990), *Improving Public Management* (London).
- Mol, N.P. (1989), 'Contract-based Management Control in Government Organizations', *International Review of Administrative Sciences*, pp. 365–379.
- (1995), 'Quality Improvement in the Dutch Department of Defence', in C. Pollitt and G. Bouckaert (eds.), *Quality Improvement of European Public Services* (London, Sage), pp. 111–120.
- Pedersen, K.J. (1977), 'A Proposed Model for Evaluation Studies', *Administrative Science Quarterly*, pp. 306–317.
- Pollitt, C. (1986), 'Beyond the Managerial Model: The Case for Broadening Performance Assessment in Government and the Public Services', *Financial Accountability & Management*, Vol 2, No. 3, pp. 155–170.
- Sorber, A. (1993), 'Performance Measurement in the Central Government Departments of the Netherlands', *Public Productivity and Management Review*, Vol. 17, No. 1, pp. 59–68.