

**PERFORMANCE MANAGEMENT AND  
ORGANIZATIONAL INTELLIGENCE:  
ADAPTING THE BALANCED SCORECARD IN  
LARVIK MUNICIPALITY**

JOSTEIN ASKIM  
UNIVERSITY OF OSLO

***ABSTRACT:** This article presents a study of how a balanced scorecard was implemented over a period of five years in four very different functional departments within Larvik municipality in Norway. The article narrates and compares the adaptation processes of the four departments, focusing on changes in their management control practices and changes in learning behavior. A surprising finding is that while management control practices of the departments varied, their learning behavior was similar. The study shows that governmental organizations from a wide range of areas of service delivery can become more active learners from adapting a performance management reform like the balanced scorecard. The article provides theoretically founded explanations of both differences and similarities in the departments' adaptation processes, and theory of organizational learning is used to inform identification of factors that can lead governmental entities into a more active learning mode.*

New Public Management (NPM) is a shorthand expression regularly used by scholars and practitioners to refer to distinctive themes, styles, and patterns of public administration that have come to the fore within the past two decades (Barzelay 2001). Following Hood (1991), we can see NPM as a body of administrative doctrines that has discredited progressive public administration's answers to administrative what-to-do questions in government. In the context of local government, performance management is the NPM doctrine that has been most widely implemented over the past twenty years (Naschold 1996; Kloot and Martin 2000, 231-232; Kleven et al. 2002, 23). This doctrine prescribes vertical decentralization and after-the-fact control mechanisms, as opposed to traditional means of public administration such as vertical centralization and before-the-fact rules and procedures (Thompson 1993; Hood 1994, 128-132). When operationally

**Direct all correspondence to:** Jostein Askim, Department of Political Science, University of Oslo, P.O. Box 1097 Blindern, NO-0317 Oslo, Norway. E-mail: [jostein.askim@stv.uio.no](mailto:jostein.askim@stv.uio.no)

specified as a tool for management of service delivery within a formal organization, performance management can take the form of a management accounting system.

A management accounting system can be defined as a system for gathering and communicating data for the ends of aiding and coordinating collective decisions in the light of the overall goals or objectives of an organization. It involves the process of identification, measurement, accumulation, analysis, preparation, interpretation, and communication of information that assists executives in fulfilling organizational objectives (Horngren, Sundem, and Stratton 1996). This article presents a case study of how the past decade's most widespread and high-profile generic management accounting system, the so-called balanced scorecard, was implemented over a period of five years in four functional departments within the Norwegian municipality of Larvik. The balanced scorecard was introduced in the early nineties by Kaplan and Norton (1992), and has dominated discussions of performance management and performance measurement since (Meyer 2002, 2).<sup>1</sup> This article raises research questions about whether and how Larvik's adaptation of the balanced scorecard affected management control and organizational learning in functional departments from very different areas of service delivery.

Studying effects on management control practices and learning behavior requires in-depth, longitudinal empirical research. That makes Larvik the best candidate research site among Norwegian municipalities, as it was among the very first Scandinavian municipalities to embark on implementing the balanced scorecard. Moreover, as the narration of the change experience will show, the implementation policy in Larvik was lax enough to give the functional departments substantial degrees of freedom in the adaptation process. This merits giving analytical interest to both similarities and differences observed across the four departments.

The first half of the article is descriptive. I will narrate the respective adaptation processes with one eye on the management control practices of the four functional departments and the other on their learning behavior. This parallel approach to studying management accounting practices is informed by March and Olsen's argument that organizational intelligence is built on two fundamental processes. The first, rational calculation, is linked to evaluation and planning, which again is linked to compliance and management control. The second process is learning from experience (March and Olsen 1976, 54). My description of the management control practices is based on a cybernetic framework, provided by Dunsire (1991). The learning behavior of the four departments is described with the aid of March and Olsen's complete cycle of organizational learning. What stands out in the description is that while the management control practices of the four departments vary substantially, their learning behavior is largely similar. Over the course of the five-year period studied, all four departments are moved into what Weick (2001) calls an active intrusiveness mode.

The article hence suggests that substantial learning effects can be gained from measuring performance in governmental organizations. This conclusion relates to a position that has emerged in the public management literature over the last few years. One recent contribution to this stance is Behn's (2003), who argues that learning from experience is one of eight purposes of performance measurement. Another contribution is from Meyer (2002), who distinguishes seven purposes of performance measurement, of which at least three correspond to learning from experience. The position might even be seen as to echo Demski and Feltham (1976), who talked about performance measurement having two purposes: facilitating decisions and

influencing decisions. The former can be identified with learning and the latter with control. In short, several authors have argued that performance measurement is well suited to support learning. But so far, empirical support of this position has been weak. This article reinforces the performance measurement and learning position with supportive in-depth and longitudinal empirical evidence.

The second half of the article seeks to provide theoretically founded explanations of differences and commonalities across the four departments' adaptation processes. I use a contingency framework based on differences between kinds of organizational technologies to set forward three hypotheses about what management control practices we might expect to develop over time in the four departments studied. These are tested against the experiences of the four departments, and two of the hypotheses are found to be of help in explaining the variance in practices. As for the learning behavior of the four departments, I find that a framework developed by Weick (2001), focusing dependency upon the environment, is helpful in explaining the similarities in behavior across the departments.

## METHODOLOGY

We begin with a brief outline of the methodology of the study. The study is based on what Yin (1994, 51) denotes as a comparative case study design, in that it compares the parallel change experiences of four functional departments within the Norwegian municipality of Larvik. My reasons for choosing Larvik as a case were outlined above. The four departments studied are the Technical Department, the Culture and Sports Department, Byskogen School, and the Work and Employment Department. This is a most similar case design (Lijphart 1975, 163); a design suited to explain differences in outcome (i.e., value of the dependent variable) among a set of very similar cases (i.e., most independent variables are equal across the cases). In my research design, the dependent variable is the departments' respective adaptations of the balanced scorecard, and especially how management control and learning behavior were affected.

In order to isolate the critical difference (i.e., the one independent variable) that explains the observed difference in outcome, the researcher must strive for similarity on as many of the control variables as possible—the independent variables that might explain the outcome. In this study, the departments' context is similar, in that they share a common national, historical, and geographical setting. Furthermore, at the start of the change experience in 1996, their administrative situation (Simon 1946) was much the same. They were all under economic strain, and all faced the challenge posed by changes in both political and administrative leadership. Most importantly, during the whole 1996-2001 period, the departments shared the same pressure from upper political and administrative management to implement the balanced scorecard management accounting system. These are all features (i.e., independent variables) that might have explained differences in outcome if they varied.

The independent variable that guided my choice of departments within Larvik was the contingency factor of organizational technology (see “Explaining Differences in Management Control Practices,” below). The four departments mentioned were neither chosen arbitrarily nor pushed forward by the CEO as his shining examples. They were selected by the researcher from the menu of sixty-five departments because they represented the desired variety in organizational technologies. The departments’ variety on this aspect provided a site for testing the hypotheses generated by the contingency theory framework.

In examining the issues at stake, I have sought to achieve a triangulation (Denzin 1978). The material used to narrate the change experiences in Larvik stems from two sources, interviews and document studies. I interviewed the chief executive officer (CEO), Einar Gaustad, the chief financial officer (CFO), Paul Hellnes, and the executive managers of each of the four departments studied. The six interviews, performed 26-27 June 2001 at the offices of the municipality of Larvik, in Larvik, Norway, were thematic, semi-directed individual interviews that lasted approximately one hour each. The interviews are stored as tape recordings, and have also been transcribed. I translated quotations from the interviews that are included in this article into English. In addition, I have studied factual accounts of the change experiences. For the whole period, 1996-2001, I have studied all annual department plans (*virksomhetsplaner*), monthly and tertiary (tri-annual) performance reports, and manager contracts from each of the four departments. None of these are available to the general public. I have also studied the publicly available annual municipal plans and municipal reports (*årsplaner og årsrapporter*).

### **CASE DESCRIPTION: MANAGEMENT CONTROL PRACTICES AND LEARNING BEHAVIOR IN LARVIK, 1996-2001**

This section describes the change experiences of the four functional departments in Larvik, with a focus on commonalities and differences in adaptations of the balanced scorecard. I will narrate the respective adaptation processes with a focus on the management control practices of the departments and on their learning behavior. I will start with management control practices, and describe the 1996-2001 change experiences in three parts—1996-1999, 1999-2000, and 2000-2001—in accordance with what my informants all agree were milestones in the Larvik change experience.

First of all, however, I will provide some background information on Larvik municipality and the four departments. This is necessary to understand the administrative situation (Simon 1946) that creates the context for the events studied.

#### **Background Information**

Larvik municipality was formed in 1988 out of the city of Larvik and the four surrounding municipalities of Hedrum, Tjølling, Stavern, and Brunlanes. Larvik is today a large municipality by Norwegian standards, with more than 40,000 inhabitants and 530 square kilometers. The administration has 2,700 employees as of 2001. Traditionally, the Norwegian Labor Party (*Arbeiderpartiet*) has governed Larvik, but since 1995 the parties in the center or at the right end of the political spectrum have been in the majority in the local council. In the period 1995-1999 the Conservative Party (*Høyre*) and the Agrarian Center Party (*Senterpartiet*) held joint power, and in the period 1999-2003 the Conservative Party shared power with the

Progressive Party (*Fremskrittspartiet*). Aud Holtskog was appointed CEO of Larvik in 1995. Gaustad took over the position early in 1999 and held it throughout the period studied and beyond.

The four functional departments studied were established as result units in 1996. The Technical Department used to be a part of the larger Technical Sector, together with the Fire Department, the Planning Department and the Electricity Department. The Technical Department is responsible for maintenance of public roads and parking spaces, including traffic safety and traffic lights; building, running, and maintaining the public water transportation and sewage systems; renovation and recycling of domestic waste; running and maintaining public parks and toilets, sports facilities, and playgrounds; and developing public residential and industrial areas. The Technical Department has ninety employees, and has been headed by Knut Hjalmar Gulliksen since 1996.

The Work and Employment Department aims to provide work or other occupation to people whose work capacity or competence is unresolved and to people who for some other reason have not made a choice of occupation (Work and Employment Department 2001, 2). The overall goal of the department is to empower people to get off benefits and into a paid occupation. Related to this, the department works to help immigrants get integrated into the community. Typical clients are immigrants without the necessary language skills, and physically challenged and mentally ill persons. The department is made up of eleven subunits, with their own managers. In total, the department has sixty employees. Since 1997, the department has been headed by Audun Auby.

The Culture and Sports Department was formerly part of the larger Culture Sector in the municipality, together with institutions like the library, some museums, and the cinema. The department is responsible for arranging and facilitating cultural activities like music, dance, theatre, literature, the arts, and sports, and other outdoor activities. The department targets both the public at large and prioritized groups like children, the elderly, immigrants, the physically disabled, and the mentally challenged. In part the department makes its own arrangements, but it also guides and economically supports relevant organizations from the nonprofit sector. Since 1996, Marit Wold has headed the department.

Byskogen School is one of the largest of the municipality's twenty-five primary schools. Eva Børven Olsen has been the principal of the school since 1995.

### **Description of Management Control Practices in Larvik**

Early in 1996 the newly elected center-conservative political coalition and the new CEO, Holtskog, agreed to start implementing a system for management control in Larvik municipality—in part as a response to demands for more economic efficiency. Management control can be defined as “the process by which managers assure that resources are obtained and used effectively and efficiently in the accomplishment of the organization's objectives” (Hofstede 1981, 193). Management control systems are the systematized means by which upper management in an organization measures, monitors and motivates the managers of its various responsibility centers (Macintosh 1985, 241).

This decision marks the beginning of the change experience studied. First, some rather drastic structural changes were made. The handful of functional sectors was broken down into sixty-nine result units. I.e., each elderly home, each cultural institution, and each school, like Byskogen, became result units—or departments—and the respective department managers were delegated powers that gave them much discretion from hierarchical interference. By mid-2001, which marked the end of the experience, a series of managerial interventions had radically altered the municipality's organizational structures and its rules and routines for expenditure planning, financial management, audit, and evaluation. Brick by brick, the municipality had built a mature performance management system.

The management control system was implemented in relation to all sixty-nine functional departments, but the implementation policy in Larvik was lax enough to give the departments substantial freedom in the adaptation process. This merits giving substantial analytical attention to both the similarities and differences I observed across the four departments studied. I will make use of two conceptual frameworks to systematically describe the management control practices that developed in the four departments in the period 1996-2001.

### *Conceptual Frameworks for Description of Management Control Practices*

The first framework is provided by Dunsire (1991), who argues that in the abstract, management control is a product of the interaction of three components. The first is a director component, which is whatever is capable of setting standards over the preferred subset of all possible states of the system. The second is a detector component, which is whatever is capable of identifying or monitoring the states of a system. The third is an effector component, which is whatever is capable of modifying behavior or altering the state of a system.

The second conceptual framework is a classification of different kinds of performance indicators (PIs). PIs relate to the director component of management control. My typology is based on two questions. The first is "What do the PIs measure?" Informed by Scott's (1992) discussion of how to measure organizational effectiveness, and Wilson's (1989) distinction between outputs and outcomes, I make a distinction between measures of outcome efficiency (outcome PIs), measures of procedural output efficiency (procedural PIs), and measures of structural output efficiency (structural PIs). The second question is, "How are the PIs used?" Carter (1991, 94-95; Carter, Klein, and Day 1992, 49-51) makes a distinction between PIs that work like dials and PIs that work like tin (can) openers. A dial provides a precise measure of, e.g., outcome, procedural, or structural effectiveness, based on a pre-set norm or standard of what makes for good or bad performance.

A dial is, e.g., a measure of the amount of bacteria in water (an outcome PI), or a pupil-per-computer ratio in a school (a structural PI). A dial is an unambiguous and unchallengeable measure; it provides a means of "stopping up excuses" on the part of subordinates (Carter 1991, 93). A tin opener is a more ambiguous PI, which by itself provides only an incomplete and inaccurate picture. Such indicators do not give answers, but prompt further interrogation and inquiry or at least discussion. An example of a tin opener PI could be a survey of client satisfaction with the level of

**TABLE 1**  
A Typology of Different Kinds of Performance Indicators (PIs)

| <i>How are the PIs used?</i> | <i>What do the PIs measure?</i> |                   |                   |
|------------------------------|---------------------------------|-------------------|-------------------|
|                              | <i>Outcome</i>                  | <i>Procedures</i> | <i>Structures</i> |
| Dials                        |                                 |                   |                   |
| Tin openers                  |                                 |                   |                   |

service they meet in a given public office. A reduction in the share of those who are very satisfied from 70 to 50 percent from one year to the next does not give any clear answers. It is simply an invitation for the responsible authorities to investigate this office, to probe, and to ask questions.

Together these two questions of what is measured and how are they used form a 2 by 3 typology of different kinds of PIs (see table 1).

In accordance with what my informants all agree are the milestones in the Larvik change experience, the four Larvik departments are narrated and compared in three parts: 1996-1999, 1999-2000, and 2000-2001. For each of the three periods I will outline which management control practices the CEO planned to implement and describe how they were actually adapted in the four departments.

*First Period: 1996-1999*

What management control practices were planned in this period? The director component was that each department was required to develop an annual budget, in which they were to identify several financial PI and some nonfinancial PIs for the coming year's activity. The detector component was a system of tertiary reporting of performance on both the financial and nonfinancial PIs. According to the 1998 annual plan, measuring and reporting of performance was in fact considered a prerequisite for the flat organizational model to function well. The main effector component was the introduction in 1998 of a rather light version of a budget flexibility system: a system for transferring the departments' over- and underconsumption of funds relative to budget from one budget year to the next. Initially, no overconsumption was transferred. The CEO thought the departments needed some time to get used to the system before it could be introduced in its full-fledged form. Furthermore, only part of the underconsumption was to be transferred. This was due to the municipality's economic difficulties at the time (Larvik Municipality 1998, 1999).

In many ways, these intended changes were adapted similarly across the four departments. They all reported tertiary on both financial and nonfinancial PIs, in accordance with the procedure decided by the CEO. The budget flexibility system was practiced similarly across the departments. Nonetheless, we can also observe that in spite of the common structures and prescriptions there were some differences in management control practices across the departments. One is whether department plans were developed, as the Culture and Sports Department did not do so in this period. A second difference relates to the number of PIs the departments reported on in their tertiary performance reports to the CEO. This varied from eleven (Work and Employment) and ten (Technical), to five (Byskogen and Culture and Sports).

*Second Period: 1999-2000*

In 1999, the economic problems in Larvik forced politicians to look for new ways of cutting costs. At this point, they brought in Gaustad as CEO to replace Holtskog. In order to further increase financial control over service production, the politicians gave Gaustad wide discretion to further develop the existing management system. Accordingly, building on the 1996-1999 foundation, Gaustad made several changes to the municipality's management system. In March 2000 he removed the remainder of the middle-line between the CEO and the departments—the three communal directors. He delegated to the department managers full authority over their respective budgets, wages, personnel issues, and service delivery. Furthermore, Gaustad altered the accounting system's format and language so that it better conformed to the textbook version of the balanced scorecard (Kaplan and Norton 1992, 1996, 2001).

The main change to the director component of management control was that all managers were required to develop annual department plans, starting in 1999 with the planning of 2000. The format of the plan was not dictated, but the departments were required to include in their plans certain economic PIs and three municipality-wide PIs for work environment. In addition, all departments were urged to develop and include in the plan a handful of PIs unique to their respective service areas, either quantified or verbal in format.

The main change in the detector component was that, from 2000, all department managers were required to write monthly and tertiary performance reports to the CEO. The monthly reports were to include quantified information on the PIs for economy and the three municipality-wide PIs for work environment. The tertiary reports were to include information on other nonfinancial performance as well, related to PIs in the department plans. The third tertiary report would also serve as the annual end-of-the-year-performance report from the department. Both the monthly and tertiary reports were to be based on standard forms.

Three important changes were made to the effector component. First, in case of deviations from the plan, both the monthly and the tertiary reports were to include information on what measures the manager had taken or was planning to take in order to get back on track. I interpret this as an effector mechanism because it would provide a strong incentive for the managers to keep their respective departments on track. Their professional pride would gear them to make efforts to avoid showing poor performance. Also, information on whether the manager needed help to get back on track was to be included. The second change was the introduction of a full-fledged budget flexibility system. From then on, both the stick and the carrot were clear. If a department overspent its budget in one year, its budget for the following year would be deducted an equivalent amount. Also, a department was allowed to keep any budget surplus for a given year, up to 5 percent of its total budget. The third effector-related change was that, from 2000, all department managers would have performance-based manager contracts with the CEO, with specified links to the PIs in their department plans. These contracts would be revised annually.



How did the departments adapt these intended changes? In many ways, Gaustad's system worked very similarly across the four departments. All departments developed plans—all but that of the Culture and Sports Department with PIs ready for assessment. All departments reported on PIs monthly and tertiary, in accordance with the procedure decided by the CEO. Furthermore, the mechanism for transferring over- and underconsumption of funds worked according to the CEO's plan in all departments. All my informants, even the managers whose departments had reduced 2000 budgets because of it, confirm that this mechanism worked well, meaning that it was practiced in what they perceived to be a fair and predictable manner. Asked how he practiced this system, the CFO said, "Where there are deviations from the budget, the reasons for it are discussed. We neither reward nor punish the departments economically if the deviations are due to things they cannot control. There is surprisingly little discussion over which is which, considering the large sums involved."<sup>2</sup>

Substantial similarities apart, at least four substantial differences in management control practices could be observed in this period. The first relates to the number of PIs used. The Technical Department, with twenty-six, had many more than the others (nine in Work and Employment and Byskogen, and five in Culture and Sports). The second difference is the kinds of PIs used. All four departments identified plenty of structural PIs in this period, but Work and Employment basically used structural dials only, while the Byskogen School used both structural dials and structural tin openers. A notable change from the first period is that the Technical Department started using structural PIs, as opposed to only outcome PIs. Byskogen stands out in being the only department not to make use of outcome dials in this period.

A third notable difference in management control practices was the role of department staff in the process leading to the department plans. There were differences between the departments in both the breadth and depth of the inclusiveness of departmental staff in the development of these plans. Finally, the use of manager contracts varied. Only two of the managers signed manager contracts in this period. The principal of Byskogen and the manager of the Culture and Sports Department did not sign, but for quite different reasons. The first did not sign for political reasons, the latter because she "forgot to."

### *Third Period: 2000-2001*

Important changes were made to the director component of management control in this last period. From 2000, department managers were required to institutionalize internal processes that involved all staff in the development of the annual department plans. A further change was the introduction of a standard macro for department plans. In this macro, one section was for PIs and another for activities, explicitly aimed at performing well on each of the PIs. Finally, managers in the so-called result support groups (*resultatsikringsgrupper*) (see next paragraph) were to quality-check each other's drafts for department plans.

The main change in the detector component was the introduction in the beginning of 2000 of the result support groups. These groups consisted of one process guide from the CEO unit and six to eight department managers from different functional sectors. Nine such groups were established. These were to meet every other month to discuss how the departments performed. At the start of each meeting, the group was to go through the monthly performance reports from each department. Managers who

were off track on any PI were required to explain what she/he had done to get back on track. Then the manager would receive help and advice from both the advisor from the CEO unit and the other department managers in the group.

Two main changes in the effector component were made. First, the CEO unit developed standard macros for the manager contracts, in which the ties to PIs in department plans were made explicit. This meant that the managers were obliged by contract to make efforts to perform well on the PIs. Second, the introduction of the above-mentioned cross-departmental result support groups meant that a strong, informal peer group review mechanism (Hood 1998) was institutionalized. This mechanism would take advantage of the managers' presumed desire to show their peers that they were doing a good job.

### *Summary of Management Control Practices in Larvik*

The narration of the Larvik change experience has shown how large, seemingly overwhelming differences across the departments' adaptations of the management reform were aligned by the CEO's stream of interventions and communicative efforts. By the end of the 1996-2001 change experience, the range of management control practices was virtually identical across the four departments studied. Macintosh (1994, 114-5) says there are four features that are helpful in distinguishing between different management control practices. These are: (i) the frequency in performance reporting, (ii) the pattern of participation in setting the department's PIs, (iii) the detail of the information provided by the PI system, and (iv) the kinds of PIs that were used (cf. the typology in table 1 above). We have seen how the first two of these features were literally aligned across the four Larvik departments by 2001.

Concerning the first point, we have seen how differences in adaptations were eradicated across the four Larvik departments as a result of the CEO's requirements and interventions. From the very beginning, all departments were obliged to follow the same reporting routines, at first tertiary, then both tertiary and monthly. As for the second point, none of the department managers were especially influential in the target-setting process in the first period studied. As Gulliksen, manager of the Technical Department said, "In the beginning I guess most of the targets [i.e., the PIs] were decided by the CEO, and then tried out in the departments subsequently. But with time we [the department managers] have been included more and more. We have looked at the targets and compared them with what we can achieve. And then we have discussions. The CEO is very good at that, I think. Sometimes I win, and sometimes he wins."

At the end of the change experience in 2001, in contrast, the CEO largely stayed out of the target-setting process and left the departments much to themselves. The author asked Auby, manager of the Work and Employment Department, about what sort of discussions he had with the CEO in the process of setting PIs: "I set the targets myself," Auby answered. "And I haven't heard any complaints yet. He [the CEO] tells us if he thinks we're not ambitious enough. 'Hot dog-targets' [*pølse mål*] he says, if he thinks we set targets that are too low. That means, you're there already. . . . But I haven't heard him complaining about my PIs, and as long as I don't hear anything I assume everything's ok."

A second aspect of the patterns of participation is the inclusion or exclusion of front-line staff in deciding the department's targets/PIs. In the first period, and to some extent in the second, the CEO allowed for substantial differences in departmental practices. But from 2000, all department managers were obliged to include all members of staff. Gulliksen, manager of the Technical Department said, "In planning 2001 we went further down in the organization than ever before. All my ninety employees participated. I let them work in groups, and asked them to make suggestions for concrete objectives. . . . We did some work to include staff in 1999 as well, planning for 2000, but not like this."

The similarities in management control practices are hence substantial. Nonetheless, substantial differences remained across the four departments in the last two of the four features Macintosh points to. One feature relates to the detail of the information provided by the PI system. The Technical Department made use of many more PIs than the other three throughout the whole period studied. With all three periods taken together, the Technical Department (with sixty-eight) stands out from the Work and Employment Department (with thirty-two) and Byskogen School (with twenty-eight), which again stand out from the Culture and Sports Department (with nineteen).

Finally, the four departments' management control practices varied when it came to what kinds of PIs were used. Related to the 3 by 2 PI typology in table 1, the big picture is that only the Technical Department identified outcome dials in the third period, none of the departments identified procedural PIs, while all four identified a vast number of structural PIs. Work and Employment basically used structural dials only, while Culture and Sports and Byskogen used both structural dials and structural tin openers. In sum, only two departments, the Technical Department and Byskogen, used a wide range of all the possible PIs. Both departments made use of both dials and tin openers through all three periods studied. (Remember, however, that the Technical Department made an even broader selection in that they used outcome as well as structural PIs.) The other two departments used dials exclusively.

### **Description of Learning Behavior in Larvik**

The rest of this section will show how the preceding section, describing management control practices, does not complete the picture of the departments' adaptations of the balanced scorecard. I will argue that we need to supplement the management control approach with a look at the learning behavior of the four departments.

This double approach is inspired and informed by March and Olsen (1976). In short and in general, I understand management in an organization as an attempt to facilitate and enhance what March and Olsen call organizational intelligence. Like individual intelligence, the authors argue, organizational intelligence is built on two fundamental processes (54). The first, rational calculation, is linked to evaluation and planning, which again is linked to compliance and management control. The second process is learning from experience. For the purposes of this article, I define organizational learning as what happens when an organization makes a relatively permanent change in its behavior on the basis of experience (Jacobsen and Thorsvik 1997, 298).

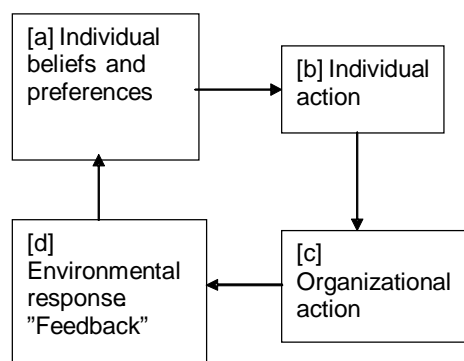
My focus on learning behavior can also be empirically justified. In his treatment of OECD public-sector reforms in the 1990s, Frieder Naschold (1996) highlights the importance of studying the process of learning in evaluations of performance management reforms. In his treatment of OECD public-sector reforms in the 1990s, Naschold notes that “[s]uccessful reform projects point to a function of the new steering model which normally remains in the background: the . . . importance of the interactive and decentralized negotiating process, [with] . . . continuous and targeted discussion and negotiation on targets, results and their conditions . . . In such a case, the new steering system is to be considered less as a [management control system] than as a suitable basis for collective and binding learning, experience and conciliation processes” (9).

### *Conceptual Frameworks for Description of Learning Behavior*

How can we frame and describe the learning behavior of the four Larvik departments? A useful starting point is March and Olsen’s (1976) complete cycle of organizational learning (see figure 1). Note that the authors do not use this cycle to illustrate that organizational learning is easy. Rather, they argue that learning is difficult because in many cases one or more of the arrows are interrupted (54-56). In describing and comparing the learning behavior of the four Larvik departments, my focus will be on how the departments—through their respective adaptations of the balanced scorecard—have attempted to build two of the arrows in the learning cycle.

The first arrow focused on is the one connecting boxes [c] and [d]. The departments’ attempts to build this arrow will be denoted scanning behavior. Scanning is a term introduced by Weick (2001, 244-246) to characterize the process of monitoring the environment and providing environmental data (feedback) to the managers of organizations.

The second arrow focused on is the one connecting boxes [d] and [a], which concerns the road from experience (feedback) to changed behavior. March and Olsen argue that in the simple situation, the individual—keen to learn—sees that he affects organizational action, which in turn affects the environment. In contrast, in situations of ambiguity we cannot trust the assumption that, e.g., negative feedback from the environment leads to changes in the beliefs and preferences of the individuals working in the organization in question. In ambiguous situations the feedback is



Source: Informed by March and Olsen (1976).

**FIGURE 1.** The Complete Cycle of Organizational Learning

unclear. Subsequent events are seen only dimly, and causal connections among events have to be inferred. It is not clear what happened or why it happened (1976, 58). This road from box [d] to [a] is arguably a long and winding one. In this study, the Larvik departments' respective attempts to build this arrow will be denoted interpretation behavior. Interpretation is a term introduced by Weick (2001, 244-246) to denote the process of translating events and developing shared understanding and conceptual schemes among members of the organization or its leaders.

A second framework that I find useful for describing learning behavior is provided by Weick (2001, 245-247). He shows how a scholar can place actual organizations on a continuum from passive to active learners based on observations of their scanning and interpretation behavior. In Weick's terminology, an organization that is an active learner is in an active intrusiveness mode. Such organizations actively scan their environment for answers concerning how well they are performing their tasks. They allocate considerable financial and human resources to continuous and systematic scanning activities. Moreover, active learners generally think creatively about how they can get feedback from different parts of their environment, and they actively engage in how to interpret whatever feedback they get. Passive organizations, on the other hand, are informal and unsystematic in their scanning and interpretation activities. Such organizations accept whatever information the environment gives them, and they interpret the environment within accepted limits (246-247).

#### *From Passive to Active Learning Behavior*

My understanding is that the implementation of the balanced scorecard in Larvik effectively helped/moved all four departments studied into a significantly more active intrusiveness mode. The description of management control practices has shown that by 2001, in contrast to the situation in 1996, a wide range of regular scanning activities took place in all four departments. All the department managers regularly obtained data/feedback from numerous sources, like monthly and tertiary assessments against PIs in department plans, biannual individual meetings with superiors, meetings with other department managers every other month, public feedback reports, and annual surveys among staff and users.

A further scanning mechanism that was not mentioned above was the introduction in 2000 of a system for so-called public feedback reports (*meldingsordningen*). The general public was urged to give the municipality reports about their satisfaction or dissatisfaction concerning any aspect of services provided. Forms for this purpose were sent to residents by mail once a year, they were made available in all municipal offices, and they could be downloaded from the Internet. Once such a form was received by the municipality it was to be distributed swiftly to the department responsible for the service in question. Within fourteen days of submittal, a complainant was to be contacted with notice of how the municipality had dealt with his or her report.

We have also seen how, by the end of the change period, all department managers engaged in extensive interpretation processes. Two important arenas had been institutionalized by the CEO to assist the department managers in interpreting the vast amount of feedback. One was individual meetings between department managers and superiors. From 2001, all managers had two annual one-to-one meetings with superiors: one between the department manager and the CEO to discuss overall

departmental performance, and one between the manager and a CEO representative to discuss the results of the user and staff surveys in particular. Olsen, principal of Byskogen primary school, emphasized the importance of these meetings: "In the case of a poor result on any of the surveys, we discuss what efforts I should be thinking about making. . . . But in these discussions, it is I as a manager who's responsible for drawing the conclusions and implementing the necessary changes. The people from the CEO unit just assist; they never take over the responsibility."

A second arena for interpretation I want to emphasize is the six annual meetings each department manager had with other department managers in the cross-departmental result support groups. These groups consisted of one process guide from the CEO unit and six to eight department managers put together from different functional sectors. All managers interviewed emphasized the interpretative aspects of these groups. Olsen said: "The first item [in the meeting] is the forthcoming monthly performance reports. We go through those quickly—are you on course [on the PIs] or not? And if you say 'no,' we ask, 'do you need help? Is there anything we can do?' . . . So you see, we lift this manager's problem in and discuss it together. And if the help from the managers in the group is not enough, the department writes in its monthly report to the CEO that it needs support."

### **EXPLAINING PATTERNS OF ADAPTATION IN LARVIK**

This section aims to provide theoretically founded explanations of important aspects of the four departments' respective adaptations of the balanced scorecard. Two questions are addressed. First, what can account for the substantial differences in management control practices? Second, why were all four departments effectively brought into a more active learning mode?

#### **Explaining Differences in Management Control Practices**

Recall that of the four features Macintosh (1994) denotes as central in distinguishing between different management control practices, only two could be aligned by the CEO's interventions. These were the pattern of participation in setting the department's PIs and the frequency in performance reporting. At the end of the period studied, substantial differences remained on the detail of the information provided by the PI system and what the PI system was used for. These differences merit an explanation that moves beyond variation in the individual wills and skills of the four department managers, which was the one given by the CEO.

The explanatory framework whose predictable power is tested in the following is the contingency approach provided by Macintosh (1985, 1994). Inspired by Wilson (1989), the contingency factor focused by Macintosh is organizational technology.<sup>3</sup> Organizational technology pertains to the nature of work activities. It is "the actions an individual performs on an object, with or without the aid of tool or mechanical devices, in order to make some change in that object" (Perrow 1970, 198). This definition stresses the organization's conversion process, which changes inputs to

outcomes. Inputs may be any sort of raw material—people, ideas, orders, paperwork, steel castings—upon which organizational skill and knowledge is brought to bear (Macintosh 1994, 112). Wilson’s point of departure for constructing his typology is as follows:

From a managerial point of view, agencies differ in two main respects: Can the activities of their operators be observed? Can the results of those activities be observed? The first factor involves outputs—what the teachers, doctors, engineers, police officers and grant-givers do on a day-to-day basis. . . . The second factor involves outcomes—how, if at all, the world changes because of the outputs. The outputs (or work) of police officers are the radio calls answered, beats walked, tickets written, accidents investigated, and arrests made. The outcomes (or results) are the changes, if any, in the level of safety, security, order, and amenity in the community (1989, 158-159).

Outputs of an agency may be hard to observe for one or both of two reasons: either what the operator does is esoteric (like an art director in an advertising firm, sitting in his chair and somehow coming up with a brilliant idea of how to sell a product), or because the operator acts out of view of the manager (like a forest ranger, looking after his forest, thousands of kilometers from his superiors) (Wilson 1989, 159). Similarly, the outcomes of an agency may be hard to observe for one or both of two reasons: either because the outcome appears after a long delay (for example, the meaningful time to measure whether a drug rehabilitation program makes an addict drug free is not the month after rehab), or because the outcome results from an unknown combination of operator behavior and other factors (for example, a pupil’s exam results reflect some mixture of pupil intelligence, parental influence, and teacher skill) (159).

Observing outputs and outcomes is neither possible or impossible; in real-life it is either easy or difficult. Nonetheless, for analytical purposes it is useful to put the two dimensions together to produce a typology of four ideal cases of organizational technologies, as illustrated in table 2.

Macintosh has extended a contingency approach very similar to Wilson’s to statements about what sort of management control practices are suitable to the different types of organizational technology. These statements are based in part on theoretical deduction and in part on induction from empirical research (1985, 247). I will use Macintosh’s framework to formulate hypotheses about the following question: What management control practices do we expect to observe developing

**TABLE 2**  
*A Typology of Organizational Technologies*

| <i>Outputs</i> | <i>Outcomes</i>       |                       |
|----------------|-----------------------|-----------------------|
|                | <i>Observable</i>     | <i>Not Observable</i> |
| Observable     | Production technology | Procedural technology |
| Not observable | Craft technology      | Coping technology     |

Source: Informed by Macintosh (1985).

over time in the relation between upper management and the four departments in Larvik?

### *Outlining Hypotheses 1-3*

An agency whose outputs and outcomes are both observable has a production technology. Wilson (1989) uses the Internal Revenue Service (IRS) as an example. The output of the IRS is the activities of clerks and auditors, and its outcome is the amount of money collected in taxes as a result of their efforts. Both are easily observed. Among the four functional departments in Larvik, the Technical Department is the one that closest resembles this ideal type.<sup>4</sup>

*Hypothesis 1.* Related to management control practice, my Macintosh-informed hypothesis would be that we would observe a close management control practice developing in the relation between upper management and the Technical Department. According to Macintosh, a close practice implies that on the director component, upper municipal management will have a great influence in setting PIs, and PIs are many and detailed, focusing on outputs as well as outcomes. As for the detector component, performance reports are detailed and frequent (1985, 248-249). This management control practice resembles what Simon et al. (1954) denote a scorecarding system—a system for close measuring and monitoring of output as well as outcome.

An agency whose outputs and outcomes are both unobservable has a coping technology. Byskogen School is the one of the four Larvik departments which closest resembles this ideal type.<sup>5</sup>

*Hypothesis 2.* Related to management control practice, my hypothesis would be that we would observe a prospects-oriented management control practice developing in the relation between upper management and Byskogen. On the director component, employees and the department managers are influential in setting targets, since they are more knowledgeable about the requirements of work, which is typically esoteric. Targets will not be detailed. On the detector component, we will find infrequent and general reporting, since greater detail would be of little help to upper management. This is a management control practice suitable for evaluation, planning, and estimation of future prospects, as well as coordination of the department's efforts with those of other departments (Macintosh 1985, 253).

In the case of craft agencies, outcomes are observable but outputs are not. Archetypal examples are advertising agencies, artists, and football teams. The outcomes of such organizations are unambiguous and readily observable from an upper manager's point of view: gross billings to clients, nice paintings, and good football, respectively. Outputs, on the other hand, are largely esoteric and hence unobservable, i.e., what upper management can decide in advance is the desired level of outcome—not how to get there. In Larvik, the Work and Employment Department and the Culture and Sports Department resemble this ideal type.<sup>6</sup>

*Hypothesis 3.* In the Work and Employment Department and the Culture and Sports Department we will observe the development of a results-oriented control practice. On the director component, targets are simple, with focus on outcomes, as close monitoring of output will not ensure efficiency or effectiveness. Department managers have great influence in setting targets. As for the detector component, control reports are frequent, and they are general in content, rather than detailed. This is a system useful in clarifying target levels and thereby helping people to do things



correctly (Macintosh 1985, 253). This resembles what Simon et al. (1954) call attention directing. If the data reveals that, e.g., the Culture and Sports Department does not achieve the desired level of outcome, they know they need to make efforts to find out what they can do differently next time around. The system helps the department direct their attention towards certain aspects of their activities.

### *Hypotheses 1-3 Revisited*

Contrary to what the Macintosh framework led us to expect, we have seen how upper management left the target setting much to the Technical Department itself. But apart from that, most of our expectations were confirmed. The PI system was detailed, focusing on outcomes as well as structural output, and many PIs were dials. And we have seen that performance reporting was frequent. In effect, we conclude that a close management control practice did indeed develop in the relation between upper management (the CEO) and the Technical Department. This means that hypothesis 1 was firmly strengthened.

Hypothesis 2 was also strengthened by the Larvik experience. Contrary to our expectations, we have seen how Byskogen's performance reporting was frequent and systematic. But most other expectations were confirmed. The target setting was left much to Byskogen's manager/principal and her employees. Furthermore, targets were not detailed. Byskogen had relatively few PIs throughout the whole period. Moreover, virtually all PIs used by Byskogen were measures of structural effectiveness, and most were tin openers. Central among these structural tin openers were surveys among staff and user groups.

What sort of control system is this? Recall that Byskogen stood out from the rest by using a wide range of all the possible PIs. The importance of this breadth was emphasized by both the CEO and the CFO. I asked the CEO, Gaustad, about what managerial use he could make of the information from, e.g., a user survey from a school:

It's clearly a diffuse measure of the performance of the principal and the school [Gaustad agreed]. Many things may influence on a user survey. But the thing is, if one school performs worse than the others on such a test year after year, we have reason to suspect that the head is not doing a good job. And remember that in addition to user satisfaction we have measures of staff satisfaction and resource consumption. . . . And hopefully we'll get indicators of pupil performances soon as well. . . . The thing is, when you put those four diffuse indicators together, you get a pretty good idea of what goes on in that school. . . . And then you can start figuring out how to make corrections.

I interpret Byskogen's experience to confirm Macintosh's expectation. Byskogen's PI system could be used by the CEO less as a tool for close measuring and monitoring than as a tool for evaluation, planning, and estimation of future prospects.

Hypothesis 3 was, however, not strengthened by the Larvik experience. In both the Work and Employment Department and in Culture and Sports, department managers and staff were influential in the target-setting process. This was what we expected to find. But this was the case in the other two departments studied as well. Also conforming to expectations is the fact that neither of the two departments issued detailed performance reports. In fact, both departments had relatively few PIs. Culture and Sports, in fact, had the fewest PIs of all four departments studied. The

reason why hypothesis 3 is still not supported is the fact that neither of the two departments used simple targets, with focus on outcomes (see hypothesis 3, above). On the contrary, both Work and Employment and Culture and Sports used structural PIs almost exclusively. Work and Employment basically used structural dials only, while Culture and Sports used both structural dials and structural tin openers. Recall that the only department to use many outcome PIs was the Technical Department.

The strengthening of hypotheses 1 and 2 and the rejection of hypothesis 3 means that Macintosh's contingency model helped us predict the management control practices of two of the four departments studied. This suggests that the model does identify factors that affect management control adaptations across different areas of service delivery. But my study also questions the robustness of the model. I will nonetheless argue that the model has shown itself helpful enough to merit further empirical research. This research should consider including aspects of Macintosh's model that were not used in my study. One candidate feature is whether the targets used are easy or hard to achieve. Moreover, future research should include more than one case within each of the four contingencies, including the category of procedural technology, which was omitted from my study altogether.

### **Why Did All Four Departments Become More Active Learners?**

The explanatory framework applied here is provided by Weick (2001). In order to explain why some organizations are in a passive and others in an active intrusiveness mode, Weick focuses on variation in dependency upon the environment (247). A heavily dependent organization needs to intrude actively into its environment. A typical example of a dependent organization is one exposed to competition. Imagine a public agency providing food to elderly residents in a municipality on a contract basis. It may be exposed to competition for next year's contract. As it is dependent on customer satisfaction for a continued contract, we would expect such an agency to be in a more active intrusiveness mode than a comparable agency that is not exposed to competition.

Related to the Larvik case, a hypothesis generated by this framework would be that the Larvik departments became more active learners in the course of the period studied because the balanced scorecard reform made them more dependent upon their respective environments.

I argue that this hypothesis was firmly strengthened by the Larvik change experience. The CEO's interventions did not expose the departments to competition, but the departments were still made considerably more dependent upon their environments. The reason for this increased dependency was the increase in transparency in the municipality. As we have seen, data about the departments' activities were obtained from numerous sources, like regular assessments of PIs in the department plans, public feedback reports, and annual surveys among staff and users. By institutionalizing routines for performance reporting, meetings with upper management, and not least the cross-departmental result support groups, the CEO exposed these data to the departments' respective environments: political and administrative upper management, users, staff, and peers (the other department managers).

This latter point, concerning making performance visible for peers, is what Hood (1998) calls institutionalizing mechanisms for peer-group review. Hood argues that peer group review mechanisms are especially useful in professional organizations. In fact, some say that what defines a professional is precisely the fact that he or she is only accountable to his or her peers (Carter, Klein, and Day 1992, 35). As a result, feedback coming from other sources—be it clients or nonprofessional managers—might easily be regarded as data of little concern rather than useful information. And the fact is that the CEO in Larvik did indeed make extensive use of peer-group review mechanisms, especially in relation to schools.

Three additional transparency-enhancing efforts were made in relation to this sector. First, the CEO introduced an additional user survey for schools. Second, the twenty-five schools in the municipality were asked to compare performance on selected indicators not only among themselves, but even with schools in neighboring municipalities. Third, the results from user surveys were given to the local newspaper, which in turn printed a best-in-test article about the twenty-five schools. From 2001, schools were also made more dependent upon another demanding actor in their environment: the parent. As the CEO described, “Some of these [PIs] obviously spur reflection, more than anything else. So I had to find a way to make them [the principals] sit down and think, ‘what does this mean.’ That’s why I’ve made the schools send letters to all parents, describing the results [of the user surveys] and stating ‘this is what we intend to do about it.’ That way you build up an expectation among the parents—an expectation they [the school managers] won’t dare to challenge.”

This illustrates my understanding of the effects of the balanced scorecard reform on the learning behavior of the departments. The transparency enhancing changes introduced by the CEO increased the departments’ perceived dependency upon their environments and made the them want to influence how their respective environments perceived their work. This is evident in the managers’ active involvement in all arenas for interpretation. The principals’ annual letters to parents is a good example.

The departments’ more active learning behavior is also evident in their ever more active scanning activities. An example here is the Technical Department. When a recycling system for domestic waste was introduced in 1999, the Technical Department received massive critique and a terrible score on the user survey. But the data from the user survey were very general, like 1-10 ratings on questions such as “How satisfied are you with the renovation services in Larvik?” This made the data difficult to interpret. Gulliksen explained how the department used the negative result: “If you ask general questions, you get general feedback. If you ask more specific questions, you get something else. I guess you can say that the general user survey was useful to us because it gave us reason to find out more. It was an indication that something had to be done.”

What the department did was to conduct a more specific user survey, with targeted questions about the renovation services and the recycling system. The feedback obtained from this survey was, in turn, used in the development of the department’s services.

## CONCLUSIONS

The article has presented a study of how a balanced scorecard, which is an operational specification of the performance management doctrine, was implemented over a period of five years in four very different functional departments within the context of Larvik municipality in Norway. These were the Technical Department, the Work and Employment Department, the Culture and Sports Department, and Byskogen Primary School. The article has narrated and compared the adaptation processes of the four departments and provided theoretically founded explanations of commonalities and differences.

The first question addressed was whether management control was facilitated by the implementation of the balanced scorecard. The narration of the departments' adaptations showed that this was indeed the case. However, we also observed how management control can take many shapes. In the relation between upper management and the Technical department, a close management practice emerged. This was a system with many detailed and unambiguous controls—a system for close measuring and monitoring of outputs as well as outcomes. In the case of Byskogen School, no such system emerged. There, measures of performance were few and ambiguous. Central features of Byskogen's PI system were measures of structural effectiveness, like data from user and staff surveys. The CEO explained how this did not imply that control was absent. When you put several ambiguous measures together, he argued, you get a pretty good idea of what goes on in that school. I.e., Byskogen's PI system facilitated prospects-oriented management control: a system that could be used by the CEO less as a tool for close measuring and monitoring than as a tool for evaluation, planning, and estimation of future prospects.

A contingency model developed by Macintosh (1985, 1994) was used in an attempt to explain the management control practices of the four departments. The model did help us to predict the adaptations and practices of two of the four. This suggests that the model does identify factors that affect management control adaptations across different areas of service delivery. And even though the study gives reason to question the robustness of the model, I argue that the model has shown itself helpful enough to merit use in further empirical research.

A surprising finding in this study is that while the management control practices of the departments varied, their learning behavior was similar. By the end of the period studied, all departments actively and regularly scanned a range of sources in their respective environments for feedback about their performances. Moreover, all departments made considerable efforts to make sense of this feedback. Innovative interpretation activities were spurred in all four departments, not least in Byskogen, underpinned by the use of arguably ambiguous data from, e.g., surveys of staff and user satisfaction. The term used in this article for such indicators of performance is tin openers. By itself, a tin opener provides only an incomplete and inaccurate picture. Such PIs do not give answers, but prompt further interrogation and inquiry, or at least discussion. We have seen how deliberations about the significance and implications of such data took place on both horizontal and vertical arenas in Larvik. The CEO institutionalized meetings between department managers and department staff, between department managers and upper management in the municipality, and last but not least, between the managers of different departments.

One lesson we can draw from this study is that performance management can thrive in areas beyond those most easily measured quantitatively. Governmental organizations from a wide range of areas of service delivery can become more active learners from adapting a performance management reform like the balanced scorecard. We have seen how even the Work and Employment Department (whose outcome is observable but whose output is not) and Byskogen School (where neither outcome nor output is observable) were cajoled into more active learning behavior. This conclusion relates to a position that has emerged in the public management literature over the last few years. Several authors have argued that performance measurement is well suited to support learning (Behn 2003; Meyer 2002; Naschold 1996). Few have provided empirical evidence to support this claim, however, and in-depth studies of local government implementations—the front-line of the welfare state—have been virtually absent.

What are the factors that can lead governmental entities into a more active learning mode? The Larvik study has identified at least two factors that can inform future, more targeted research into this more practical question. One is the importance of measuring a wide range of the activities and performance of the entity in question. A typology of different kinds of performance indicators was used to illustrate this width in this article (see table 1). A second potentially important factor is the institutionalizing of various and supplementing arenas for interpretation and deliberation, and hence learning from measures of performance.

**ACKNOWLEDGMENTS:** The author is grateful for comments and advice from Harald Baldersheim, Michael Barzelay, Niels Ejersbo, Mark Thatcher, Fred Thompson, three anonymous reviewers, and the participants in the 11th Nordic Conference on Local Government Research, University of Southern Denmark, Odense, 29 November-1 December 2002. He is also grateful to Einar Gaustad and the other interviewees in Larvik municipality, Norway.

## NOTES

A previous version of this article was presented at the 11th Nordic Conference on Local Government Research at the University of Southern Denmark, Odense, 29 November-1 December 2002.

1. Note that at the time of writing (May 2004), Amazon holds more than twenty hardcover books with balanced scorecard in the title. A search in ISI Web of Knowledge returns eighty-six articles with balanced scorecard in the title.

2. Note, however, that for the whole period 1996-2001 there were only two instances where the budget transfer mechanism was not practiced as a dial in Larvik, meaning that over- or underconsumption was not directly transferred to the next year's budget. One example is from the Work and Employment Department, which in 2000 had an overconsumption of NOK 350,732 that was not transferred to the 2001 budget. The CFO explained that this was a reward for the department's hard work in 2000, which resulted in a record rate in getting benefit claimants into paid occupations, which in turn saved the municipality NOK 5 million in benefit payments.

3. Note that using a contingency approach is not a novel way of studying management accounting practices. It corresponds to focusing on differences between organizations, which is one of three different contingency approaches that scholars have used to analyze the use and purposes of performance measurement, as identified by Noordegraaf and Abma (2003). A second type of contingency framework puts emphasis on differences in (production) processes. An example of this is Hofstede (1981), who distinguishes, e.g., repetitive from

nonrepetitive processes. A third contingency approach focuses on differences in substantial issues. Examples are crisp vs. fuzzy issues (Lerner and Wanat 1983) and well known vs. little known and contested vs. not contested issues (Bowker and Star 1999; Noordegraaf and Amba 2003).

4. Recall that the Technical Department is responsible for public roads and parking spaces, public water transportation and sewage systems, domestic waste, public parks and toilets, sports facilities and playgrounds, and for developing public residential and industrial areas.

5. Recall that outputs of an agency may be hard to observe either because of what the operator does or because the operator acts out of view of the manager. The output of a teacher at Byskogen is unobservable for both reasons. His or her output is esoteric because no manager can predict exactly how a teacher can succeed in explaining algebra to a specific pupil. And, a teacher behind the closed classroom door is well out of view of superiors. The outcomes of a school are also unobservable, because the outcome is the result of an unknown combination of the teacher's or the school's behavior and other factors. Alternatively, if one takes the view that the outcome of schooling is something like wise, tolerant, and well-adjusted citizens rather than numeric exam results, the outcome obviously appears after a long delay.

6. Recall that the outcome of the Work and Employment Department is to provide work or other occupation to people whose work capacity or competence is unresolved and to people who for some other reason have not made a choice of occupation. Typical clients are immigrants without the necessary language skills, the physically challenged, and the mentally ill. The outcome is to get people off benefits and into a paid occupation. Whether or not this happens is fairly easy to observe. The department's output is harder to observe. Much of the work consists of face-to-face interaction between councilors and clients, where the councilor aims to increase the client's self-esteem and bring him/her out of passivity. Exactly how a councilor manages to empower and motivate a specific client is hard to observe.

Similarly, the outcome of the work of the Culture and Sports Department is cultural events like music, dance, theatre, literature, and arts, as well as sports and other outdoor activities. The department targets both the public at large and prioritized groups like children, the elderly, immigrants, the physically disabled, and the mentally challenged. Sometimes the department is directly responsible as the arranger of activities, but most often the department has a more intermediate role. The department advises, guides, and economically supports cultural and sports organizations from the nonprofit sector in the municipality. The department also invites external actors like the national theatre (*Riksteateret*) to perform in Larvik. Whether or not popular cultural and sports activities take place in Larvik is fairly easy to observe. The output of the Culture and Sports Department is harder to observe, especially in the areas where the department has an intermediate role. Then, most of the work is unobservable, like network building within and outside the municipality, and inspiring and empowering local organizations.

## REFERENCES

- Barzelay, M. 2001. *The New Public Management: Improving Research and Policy Dialogue*. Berkeley: University of California Press.
- Behn, R.B. 2003. "Why Measure Performance? Different Purposes Require Different Measures." *Public Administration Review* 63:586-606.
- Bowker, G.C., and S.L. Star. 1999. *Sorting Things Out*. Cambridge, Mass.: MIT Press.
- Carter, N. 1991. "Learning to Measure Performance: The Use of Performance Indicators in Organizations." *Public Administration* 69:85-101.

- Carter, N., R. Klein, and P. Day. 1992. *How Organizations Measure Success. The Use of Performance Indicators in Government*. London: Routledge.
- Demski, J.S., and G.A. Feltham. 1976. *Cost Determination: A Conceptual Approach*. Ames: The Iowa State University Press.
- Denzin, N.K. 1978. *Sociological Methods: A Sourcebook*. New York: McGraw-Hill.
- Dunsire, A. 1991. "A Cybernetic View of Guidance, Control and Evaluation in the Public Sector." Pp. 325-45 in F.X. Kaufmann, ed., *The Public Sector*. Berlin: de Gruyter.
- Hofstede, G. 1981. "Management Control of Public and Not-for-Profit Activities." *Accounting, Organizations and Society* 6:193-211.
- Hood, C. 1991. "A Public Management for all Seasons?" *Public Administration* 69:3-19.
- . 1994. *Explaining Economic Policy Reversals*. Buckingham, UK: Open University Press.
- . 1998. *The Art of the State*. Oxford, UK: Clarendon.
- Horngren C.T., G.L. Sundem, and W.O. Stratton. 1996. *Introduction to Management Accounting*. London: Prentice-Hall.
- Jacobsen, D.I., and J. Thorsvik. 1997. *Hvordan organisasjoner fungerer: Innføring i organisasjon og ledelse* (How Organizations Work: Introduction to Organization and Leadership). Bergen, Norway: Fagbokforlaget.
- Kaplan, R.S., and D.P. Norton. 1992. "The Balanced Scorecard." *Harvard Business Review* 70:71-8.
- . 1996. *The Balanced Scorecard: Translating Strategy into Action*. Boston: Harvard Business School Press.
- . 2001. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston: Harvard Business School Press.
- Kleven, T., ed., E.H. Allern, N. Ejersbo, P. Hansen, S. Montin, J. Mørch, and O. Rieper. 2002. *Målstyring i Skandinaviske kommuner. Seks casestudier* (Management by Objectives in Scandinavian Municipalities: Six Case Studies). NIBR-rapport 2002:15. Oslo: Norsk institutt for by- og regionforskning.
- Kloot, L., and J. Martin. 2000. "Strategic Performance Management: A Balanced Approach to Performance Management Issues in Local Government." *Management Accounting Research* 11:231-51.
- Larvik Municipality. 1998. *Årsrapport 1998* (Annual Report 1998). Larvik, Norway: Larvik Municipality.
- . 1999. *Årsrapport 1999* (Annual Report 1999). Larvik, Norway: Larvik Municipality.
- Lerner, A.W., and J. Wanat. 1983. "Fuzziness and Bureaucracy." *Public Administration Review* 43:500-9.
- Lijphart, A. 1975. "The Comparative Cases Strategy in Comparative Research." *Political Studies* 8:158-77.
- Macintosh, N.B. 1985. *The Social Software of Accounting and Information Systems*. London: John Wiley and Sons.
- . 1994. *Management Accounting and Control Systems*. New York: Chichester.
- March, J., and J.P. Olsen. 1976. *Ambiguity and Choice in Organizations*. Bergen, Norway: Universitetsforlaget.
- Meyer, M.W. 2002. *Rethinking Performance Measurement: Beyond the Balanced Scorecard*. Cambridge, UK: Cambridge University Press.
- Naschold, F. 1996. *New Frontiers in Public Sector Management. Trends and Issues in State and Local Government in Europe*. Berlin: de Gruyter.
- Noordegraf, M., and T. Abma. 2003. "Management by Measurement? Public Management Practices amidst Ambiguity." *Public Administration* 8:853-71.
- Perrow, C. 1970. *Organizational Analysis: A Sociological View*. London: Tavistock.

- Scott, W.R. 1992. *Organizations: Rational, Natural and Open Systems*. Upper Saddle River, N.J.: Prentice Hall.
- Simon, H.A. 1946. "The Proverbs of Administration." *Public Administration Review* 6:53-67.
- Simon, H., H. Guetzkow, G. Kozmetsky, and G. Tyndall. 1954. *Centralization vs. Decentralization in Organizing the Controller's Department*. New York: Controllership Foundation Inc.
- Thompson, F. 1993. "Matching Responsibilities with Tactics: Administrative Controls in Modern Government." *Public Administration Review* 53:303-18.
- Weick, K.E. 2001. *Making Sense of the Organisation*. Oxford, UK: Blackwell.
- Wilson, J.Q. 1989. *Bureaucracy*. New York: Basic Books.
- Work and Employment Department (*Arbeidslivsenheten*). 2001. *Virksomhetsplan 2001* (Department Plan 2001). Larvik, Norway: Larvik Municipality.
- Yin, R.K. 1994. *Case Study Research: Design and Methods*. Thousand Oaks, Cal.: Sage.



