MEASURES VS. ACTIONS: THE BALANCED SCORECARD IN SWEDISH LAW ENFORCEMENT*

Salvador Carmona

Anders Grönlund

Instituto de Empresa
Dept. Accounting Finance &
Management Control
C / María de Molina 12, 5ª,
28006 – Madrid, España
scarmona@profesor.ie.edu

Stockholm University

Abstract

Studies of organizational performance have overwhelmingly relied on evidence gathered from private sector firms. Nevertheless, the past several years have witnessed increasing interest in enhancing effectiveness and efficiency in the public sector, in turn generating considerable investment in the deployment of performance metrics in such settings. Though extant evidence provides many perceptive insights into the specifics of performance frameworks in public sector organizations, little is known about the measurement of organizational performance in police work. Our investigation drew upon the deployment of the balanced scorecard in Swedish Law Enforcement, an organization that long ago implemented the new paradigm of policing, which consisted in enhancing the quality of urban life on the mere making of arrests. Results from this investigation concur with other studies indicating that public sector organizations tend to assume a stakeholder perspective on performance measurement. In particular, Swedish Law Enforcement developed a set of measures of external success and internal performance that addressed present, past, and future time dimensions. Implementation of the balanced scorecard in police work, though, raised some problems. Our study details concerns about the aggregation of non-financial performance measures. More importantly, some crucial areas in the new concept of policing (such as community policing) were neglected by the system. Conversely, the system focused on monitoring some easy-to-measure indicators that provided a traditional view of police work while some crucial areas of policing were not measured. This focus ultimately lessened the operational potential of the balanced scorecard system. Our study also puts forward some suggestions for future research in this area.

Introduction

Investigations of organizational performance have increased during the past several years (Neely, 1999). The idea that "performance measurement matters" has resulted in the proliferation of various frameworks of organizational performance: these include Performance Measurement (Lynch and Cross, 1991), the Results and Determinants Framework (Fitzgerald *et al.,* 1991), Performance Measurement for World Class Manufacturing (Maskell, 1991), the Balanced Scorecard (Kaplan and Norton, 1992, 1996), the Cambridge Performance Measurement Design Process (Neely et al., 1996, 1997), the Reference Model of Integrated Performance Measurement System (Bititci *et al.*, 2000a), and the Performance Prism (Neely, Adams and Kennerley, 2002), to name a few. At the same time, both public sector and non-profit organizations have experienced increasing demands for more effective decision-making and more efficient management of resources (Brunsson, 1994; Brignall and Modell, 2000). Pressures from constituents have brought about the deployment of market-based control models in non-profit and governmental organizations (Kaplan and Norton, 2001): health care agencies (Van Peursem, Pratt and Lawrence, 1995) and local governments (Palmer, 1993). In spite of our increasing understanding of performance measurements within the public sector, little is known about the adoption patterns of performance metrics in policing groups. A study that addressed police work would expand extant knowledge about the universe of settings that enforce performance measurement systems; it would also advance understanding about the design of performance frameworks in governmental organizations.

Our study examines the deployment of the balanced scorecard, a performance measurement system that enables managerial decision making by aligning performance indicators with the goals and strategies of the firm (Lipe and Salterio, 2000: 284). The balanced scorecard has attracted considerable interest in the realms of practice and research. For example, Silk (1998) reports that 60 percent of Fortune 1000 firms have experimented with the balanced scorecard. Further, Kald and Nilsson (2000) show that 27 percent of major Scandinavian companies have implemented this performance measurement framework. In a similar vein, Atkinson and Epstein (2000b: 2) echo the conclusions of a study by Walker Information which reports that 59 percent of Canadian executives claim familiarity with the terms "balanced scorecard" or "balanced measurement system" (Walker Information, 1998: 4). Lastly, research interest in the balanced scorecard is reflected in the contention by Atkinson *et al.* (1997a: 94) that investigation of such performance measurement frameworks constitutes one of the most significant developments in management control and, thus, deserves intense research attention.

Empirical evidence supporting this study was gathered from the implementation of the balanced scorecard in police work in Sweden. In 1997, the Swedish National Police Board, the governing board of Swedish Law Enforcement, considered the deployment of a program to assess organizational performance in a number of districts across the country. In this paper, we report on the antecedents of that program as well as its implementation. Our observation period goes from 1998 through 2000, that is, until the project consistently reported some results.

Our investigation may be of interest for several reasons. First, this paper addresses the design and implementation of the balanced scorecard in a public sector organization. Admittedly, most academic discussions on performance measurement issues focus on

private sector firms (Atkinson and Epstein, 2000b; Kaplan and Norton, 2001: 97). Ittner and Larcker (1998: 229), however, note that "recent efforts to 'reinvent' the government have emphasized the important role performance measurement systems can play in improving the efficiency and effectiveness of government operations", and this has resulted in a special concern within the public sector about performance indicators. Such concern is exemplified by governmental investment of hundreds of millions of dollars in performance measurement systems that attempt to make governments' decisions more effective and their operations more efficient (Atkinson, Waterhouse and Wells, 1997b: 26).

The extant literature suggests at least two reasons for public sector concerns about performance measurement. First, in contrast with their use in profit-seeking firms, financial measures in the public sector do little to ascertain whether a government agency is delivering on its mission (Kaplan and Norton, 2001: 98-99). Second, public sector agencies are characterized by the need to translate political goals, primary goals, into concrete actions, secondary goals, or scorecards (e.g., Atkinson et al., 1997b). We concur with Kaplan and Norton's contention (2001) that this idiosyncratic complexity of public sector organizations introduces considerable potential for revising the architecture of the balanced scorecard.

A second reason our investigation may be of interest is that, in spite of the growing literature on performance measurement systems in government agencies, not much is known about the deployment of the balanced scorecard in policing groups. By examining police work, we expect to advance knowledge about performance metrics in an organization that regards concealment and enforcement as fundamental principles. This sense of secrecy and concomitant weak practices of accountability lead to limited expectations of both police work and its performance measurement, which in turn result in the failure of the public to recognize the important contributions that police make to the quality of urban life (Moore and Poethig, 1998). Accordingly, our study may cast light on performance measurement systems in other organizations characterized by a strict sense of secrecy, discipline, and hierarchy.

Lastly, it has been argued that stakeholders' demands exhibit differences across countries and time periods (Olson, Guthrie and Humphrey, 1998), and a number of commentators suggest that there are international differences in the deployment of performance measurement systems in public sector organizations (Hood, 1995; Brignall and Modell, 2000). The Swedish setting, we argue, characterized by its traditional, *Scandinavian,* focus on employee and social issues (Atkinson and Epstein, 2000a: 24), may provide many perceptive insights into performance measurement systems in public sector organizations.

The remainder of this paper is structured as follows: The paper initially describes the foundations of the balanced scorecard framework. This is followed by a description of the wider context of police work in Sweden and the specifics of the design and implementation processes.

Finally, we analyze our results, outline the shortcomings of this investigation, and put forward some suggestions for future research in this area.

Methods

In 1997, one of the members of the research team received word that the Swedish National Police Board was interested in deploying a program of performance measurements to assess and assist police work through balanced scorecard. The researcher, who was struck by the lack of empirical evidence on performance measurement systems in the public sector at that time, considered that available proposals on balanced scorecards' dimensions had to be redefined to fit in the specifics of police work. Additionally, the researcher questioned himself about the functioning of performance measurement systems in policing. With said research questions in mind and in the context of the active relationships between the public sector and universities in Sweden, the researcher asked the National Police Board for permission to participate in the project as an external observer (Berg, 1998). The National Police Board granted the researcher access to the project, which involved unrestricted interaction with staff as well as unlimited access to internal documents on this project. The National Police Board also provided the researcher with free access to its library resources, which contain valuable information about the contexts and specifics of police work. The researcher's involvement in the project went from 1998 through 2000, that is, until the balanced scorecard program concluded its implementation stage. During the project, the researcher was accompanied by two research assistants as well as by a doctoral student who was experienced in policing and had an interest in investigating management of police work.

Field research comprised several qualitative methods to get a sense of the structures and processes that characterized the implementation of balanced scorecards at Swedish Law Enforcement (Rossman and Rallis, 1998). To be immersed in the setting, the researcher developed a participant observation approach to the project, which consisted of attendance at planning and control meetings along with informal conversations with police staff (Marshall and Rossman, 1999). Under this approach, the researcher witnessed discussions between police officers about their hopes and concerns regarding the implementation of the balanced scorecard in Swedish Law Enforcement. Furthermore, the researcher also engaged in informal conversations that provided him with valuable information about the particulars of the project. During this stage of the project, written notes were taken but conversations were not taped. Additionally, the researcher reviewed extensive internal memoranda as well as public documents that are inherent to police work. Once the implementation of the balanced scorecard project was well under way, the researcher conducted a series of ten semi-structured interviews with police officers to unfold the participants' perspectives on the project. Interviews were held with people in positions at all staff levels that participated in the project. All interviews were taped and full written transcriptions were made. Coding of interviews and other written notes focused on aspects of actual use of balanced scorecard indicators in practice (e.g., decision-making, control), attainment of goals, areas of policing that had experienced substantial improvement as a consequence of the implementation of balanced scorecard, areas that could still benefit from using balanced scorecard figures, and overall satisfaction with the project. In concordance with Wilson (1977: 255), this combination of qualitative research methods was intended to unearth archival records and documents and verbal communications, both among participants and between them and the researcher.

The balanced scorecard

Financial performance measures have been severely criticized for their historical focus and short-term emphasis (Kaplan, 1983). Such criticisms have motivated the emergence of a number of non-financial performance metrics in the past several years (e.g., quality and inventory indicators). As noted by Medory and Steeple (2000: 521), "Non-financial measures are more timely than financial ones; the measures are very measurable and precise; the measures are meaningful to the workforce so aiding continual improvement."

Implementation of non-financial indicators, though, presents problems: the extent to which such measures may be aggregated is debatable (Dervitsiotis, 1997; Bond, 1999). Further, the link between non-financial and financial indicators is unclear (Kaplan and Norton, 1992; Dervitsiotis, 1997). To resolve these problems, Kaplan and Norton (1992, 1996) developed the notion of the balanced scorecard, which is regarded as a complement to, rather than a replacement of, financial measures. Instead of consisting of a mere collection of financial and non-financial measures, the scorecard comprises four interrelated "perspectives": financial, customer, internal business, and innovation and learning (see Figure 1). The perspectives, it is argued, are composed of leading and lagging indicators. Whereas the former drive performance (e.g., measures of internal business, innovation and learning), the latter benefit from performance drivers (e.g., indicators addressing financial and customer issues). In short, the balanced scorecard constitutes a systematic attempt to measure the relationship between results and operating activities as well as a powerful instrument to communicate a firm's goals and objectives to operating managers (Atkinson and Epstein, 2000a: 24, 27; Norreklit, 2000: 20).

Figure 1:
The Balanced Scorecard Framework
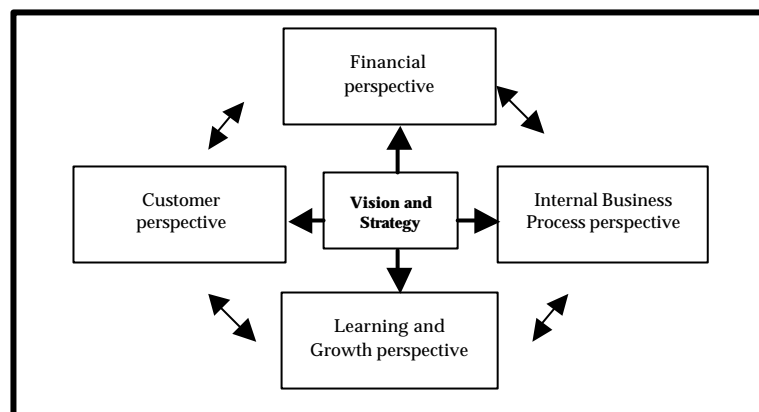
The Stakeholder Approach to the Balanced Scorecard.

Since its inception in 1992, the notion of the balanced scorecard has experienced revisions and extensions. One such extension is of particular interest to this paper—the stakeholder approach to the balanced scorecard. This approach identifies the major constituents of the firm and establishes scorecards that align the goals of the organization with the interests of its stakeholders (Atkinson *et al.*, 1997b). In this manner, a firm may map demands arising from stakeholders (e.g., shareholders, customers, employees, suppliers, and the general public) and develop a scorecard of measures that target such pressures. For example, Epstein and Birchard (1999: 149-154) identify four categories of stakeholders. The first is shareholders, to whom a firm should report on the extent to which its strategy is financially successful. The second category is customers, who must be convinced that they are receiving the value they expect. Third is employees, whose contribution is crucial for the firm to fulfil its strategy; these stakeholders should receive data on the extent to which the firm delivers value to them in return for their contribution to organizational goals. The fourth category is the communities, which also help firms to attain their goals but should also be informed on the extent to which organizations deliver value in return. In short, by deploying a stakeholder approach to the balanced scorecard, a firm communicates its priorities to constituents, which in turn are the main providers of a firm's resources and, ultimately, the underpinnings of its long-term survival.

Interestingly, the stakeholder scorecard has been implemented in organizations within both the private and non-profit sectors. For example, the Bank of Montreal established its scorecards as a function of its four major stakeholders: shareholders, customers, employees, and communities (*e.g.,* Atkinson and Epstein, 2000 a,b). For shareholders, primary measures comprised return on common shareholders' investment and secondary measures such as revenue growth. Customer satisfaction, in turn, exemplified a primary measure for customer stakeholders, whereas customer surveys for different products were deployed to test for secondary measures. Commitment, competence and employee productivity constituted primary measures for this stakeholder, which in turn resulted in secondary measures gathered from an employee opinion survey. Lastly, the community was primarily measured through the public image of the Bank of Montreal, which ultimately involved different external surveys to account for this. Interestingly, the Bank of Montreal focused on shareholders (40%) over customers (30%), employees (20%) and communities (10%). In contrast, Epstein and Birchard (1999: 149) report that the five categories of stakeholders that comprised the balanced scorecard of Eastman Chemical of Kingsport were given equal priority.

The Ontario Hospital Association provides another example of the stakeholder approach to balanced scorecard. This organization implemented said performance measurement system by encompassing the perspectives of financial performance, patient satisfaction, clinical use and outcomes, and system integration and change (*e.g.,* Atkinson and Epstein, 2000 a,b).

Police work

Policing has witnessed a changing paradigm over recent decades, shifting from the narrow perspective of reducing crime through making arrests to the wider view of strengthening its role in community life (Moore and Poethig, 1998). The former is

considered reactive and primarily reliant on calls to emergency numbers; it is associated with policing that is considerably inefficient for its usually late arrival at the scene (Sparrow, Kennedy and Moore, 1992). By contrast, the latter paradigm stresses the need to complement the traditional system by enhancing "community policing" or "neighbourhood policing"—meaning that law officers seek insights from citizens. To put it briefly, traditional attention to arrests, clearance rates, and speed of responses to calls for police service is being complemented by increasing emphasis on crime prevention, fear reduction and, eventually, an enhancement of the role of the police in raising the quality of urban life (Moore, Spelman, Estrich and McGillis, 1985).

Sweden a long ago implemented the the new policing paradigm. In 1959, a Royal Circular enacted the new role of "county councils and local police authorities on the subject of crime prevention work by the police" (Swedish Code of Statutes, 1959). Accordingly, the police were required by § 2, s 1 of the Police Act to prevent crime and other disturbances of the peace or safety of the general public. Further, section 2 stated that police should preserve peace and general security, prevent disturbances of the peace, and intervene when such breaches arise. This approach to crime prevention was reinforced by recent regulation. For example, the Government Bill 1989/1990: 199 established that "a greater proportion of total police resources should be allocated for crime prevention which, in turn, would necessitate a change in work methods." In a similar vein, the 1992/1993 Budget Bill stated that "crime prevention operations should form a natural component in policing, …[crime prevention] will require a problem-oriented approach to policing." Lastly, the 1994/1995 Budget Bill stated that "the concerted effort on the road towards problem-oriented neighbourhood policing must … be implemented within the framework of police training at the National Police College where the new police-in-the-making should receive training adapted to meet the requirements of the new methods of work." In short, these regulations indicate political, legal, and financial support for the role of police in crime prevention.

The approach of the Swedish government to crime prevention has implications for the organization of police work. At the local, district level, policing comprises three main activities: community policing, maintaining public order, and crime investigation. Community policing tackles crime prevention by instructing those without a criminal record to refrain from crime. In the main, community policing involves promotion of police/public partnership, and this results in actions such as providing educational guidance to school children who bear a high risk of becoming "permanent customers of the judicial system" and, similarly, encouraging those who have committed crimes not to repeat that behaviour. Common problems handled by community policing are peddling alcohol to youth, drug offences (by youths and adults), and drunk driving.
The maintenance of public order requires the authorities to attempt to prevent crime and other disturbances of the peace as well as to take action when a crime is committed. In the main, maintaining public order involves protection, advice, and assistance to the public through specific actions such as supervision of places frequented by criminal and other anti-social individuals, monitoring of habitual criminals, and policing of special events such as demonstrations, rallies, and state visits.

Lastly, crime investigation exemplifies the reactive type of work that people often perceive as typical of police work. When an indictable offence has been committed, the police open an investigation with the purpose of identifying the perpetrator. The officer in charge is responsible for making a case to be presented in court. Interestingly, the

heavy burden for investigators is reflected in the backlog or investigation balance, which measures the difference between registered and unsolved cases.

Admittedly, the new role of the police poses some problems in setting secondary goals and management of operations. For example, police are required by the public to respond to calls for service in situations ranging from medical problems to crime emergencies. It is not unusual, however, for citizens to call police to mediate minor disputes or to find ways to get into locked apartments and cars. Serious crime calls obviously take priority. However, insofar as police stations are among the few government agencies that are open 24 hours a day, seven days a week, many late night and weekend emergencies fall into the hands of the police, regardless of the importance of the problem within the priorities of the law enforcement organization (Sparrow *et al.*, 1992). This situation, we contend, outlines an apparent contradiction between what police should do and what they actually do. As noted by Moore and Poethig (1998), the focus of police work on reducing serious crime neglects the role of police in managing disorder in public places, reducing fear of crime, controlling traffic and crowds, and providing services in various emergency situations. Though the *missing* dimensions of police work are increasingly important and have a significant impact on the quality of urban life, they are difficult to measure. Therefore, metrics of police work typically concentrate on easy-to-measure dimensions such as number of arrests, response times, and backlog balances of clearance rates (Dilulio *et al.*, 1994), which in turn make officers appear competent in terms of the indicators employed, but which narrow their focus to the traditional model of policing.

The balanced scorecard in swedish law enforcement

Designing the Balanced Scorecard

The Swedish National Police Board (NPB) constitutes the top-level authority for police service. Its role primarily consists in providing advice, assistance, and co-ordination for the efforts at regional and local police levels. The NPB is organized into 21 decentralized county police regions. The regions have a county board that enjoys a significant degree of autonomy in activities such as planning and allocation of resources. The county police superiors typically hold a degree in law and have limited personal experience in ordinary police work. Counties are, in turn, organized into districts.

Swedish Law Enforcement has an information system that produces statistics on more than one hundred indicators. Measures are reported to the NPB on a monthly basis from local districts, via county authorities. Since the budget is set annually, monthly reports focus on non-financial indicators like crime rates, number of hours of foot patrol, and number of alcohol tests administered. Interestingly, feedback to local police from the central authorities stems from variation reports that are discussed in monthly review meetings. Nevertheless, this reporting system was deemed unsatisfactory by all layers of the police chain of command: the NPB, the counties, and the districts. A middle-level officer said, "We've got better data from reading the tabloids."

In 1998, six county police authorities set up a balanced scorecard project in co-ordination with the NPB. Motivation for the implementation of the balanced scorecard in Swedish Law Enforcement vis-à-vis other performance measurement systems was the recent, unsuccessful experience of the police organization with another performance measurement project. This prompted the NPB to consider the implementation of the balanced scorecard, a performance measurement system that had a reputation for success and that was endorsed by a leading higher education institution, the Harvard Business School. The scorecard aimed to complement the formal system on crime statistics and to enhance planning and control systems within Swedish Law Enforcement. Furthermore, the scorecard raised expectations as a communication device and, consequently, was seen as a tool in implementing the performance framework as well as improving organizational communication, both vertically and horizontally. It was thought also that discussions between staff and management on operational goals brought about higher performance results than when goals were set unilaterally by management.

The driving idea of the balanced scorecard model in Swedish Law Enforcement was that performance indicators referring to goals/success factors had to be supplemented with information about public opinion concerning local problems and levels of crime victimization or fear of crime. To tackle this objective, the model attempted to develop a triple perspective of past/present/future indicators. The past was captured through historical ratios; the present was reflected through measurement of resources, such as governmental grants, budget, and number of staff; the future was incorporated into the system through forecasts of resources, assessments of staff opportunities and challenges, and public opinion about the work of the police.

The Swedish Law Enforcement scorecard comprised four perspectives: *success, staff, citizen,* and *resources*. The staff perspective addressed personnel as stakeholders in police work, whereas the resources perspective encompassed the government as the main supplier of funding, material and human resources. The citizen perspective accounted for the reaction of the public towards policing (see Figure 2). Indicators were reported through a traffic light colour system, where green depicted a fine performance; yellow, an acceptable situation; and red, a poor performance.

Figure 2:
The Balanced Scorecard in Swedish Law Enforcement
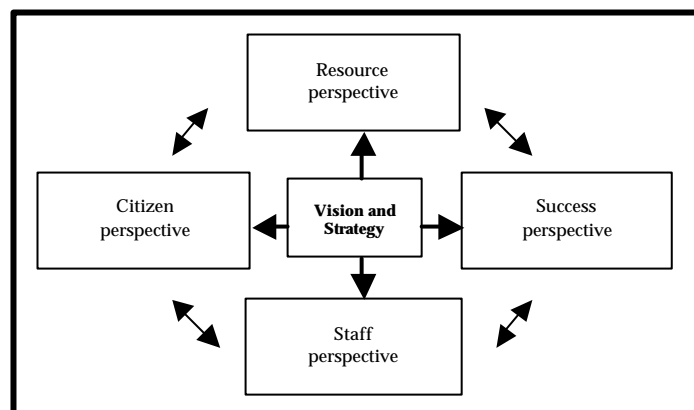
The *success perspective* was regarded as "the most crucial strategic dimension within the balanced scorecard model." When developing this perspective, police authorities started by asking three questions:

- What do we want to achieve?

- When can we achieve these goals?

- How do we measure our progress?

Drawing on discussions with staff, police authorities determined quantitative indicators of success, which included measures of violent crime, drug crime, road traffic, theft and criminal damage, public order, youth crime, crime victims, community policing, and crime investigation. Measures referred to concrete priorities that established how a unit should operate in the short term to achieve its long-term goals and thus consisted of clearly defined and measurable factors collected on a monthly basis. Exhibit 1 provides an example of the breakdown of four success factors (e.g., road traffic, public order, drug crime and crime investigation) into formulation of goals, indicators, measurement and methodology. As shown in Exhibit 1, goals were formulated in a general manner (e.g., for public order, increase safety in public areas during weekends), which in turn resulted in indicators like number of hours of police patrol in public areas during weekends. Each indicator was ultimately measured as green/yellow/red depending on the number of hours of policing at public areas, whose method involved foot and car patrol.

Exhibit 1:
Success Factors in the Balanced Scorecard of Swedish Law Enforcement

Success factors	Formulation of goals	Indicators	Measurement	Methodology
Road Traffic From the Order & Safety Department.	Discover drunk-drivers through alcohol and drug tests	Number breathalysers/month. Number of Rf[1]/month.	Number of breathalysers ≥ 200 = Green 100-199 = Yellow < 100 = Red	Directed controls around pub-areas. Controls at predetermined places. Directed controls around major parties.
Public Order From the Order & Safety Department.	Increase the safety at the pub-related centre areas during Friday and Saturday nights.	Number of resource hours/month. Number of Lob[2], Pl13[3], arrested/month and order-position.	20 h = Green 13-19 h = Yellow 0-12 h = Red 1-2 = Green 3 = Yellow 4-5 = Red	Through being at the area for point duty and foot-patrol and car-patrol. Intervention in case of need.
Drug Crime From the Local Police Department.	Directed operation against drug crime. Discover new young people that are addicted to drugs.	Number of resource hours. Number of drug crimes. New drug addicts should be measured in the number of memorandums to the welfare authorities. All numbers are presented per month.	24 h & ≥ 1 drug crime = Green 24 h = Yellow < 24 h = Red	During at least one night per month a directed operation against drugs is to take place. At least 4 policemen should work on these occasions.
Crime Investigation From the Criminal Department.	Through group subdivision of processing urgent cases. Inner search; increase the percent of solved errands, cases with unknown perpetrator.	Number of cases/month. Number of cases with a searched suspect. Number of cases with a search effort of 500 h/year.	< 1 200 cases = Green 1 200 – 1 300 = Yellow > 1 300 = Red 25% = Green 10 – 24% = Yellow 0 – 9% = Red	A 'balance-scorecard group' is set up to render possible more efficient case handling. Have a special inner search, registration and contact with Kut[4].

---

[1] Swedish law concerning arresting a drunk driver
[2] Swedish law concerning when a drunk is taken into custody
[3] Swedish law concerning troublesome behaviour
[4] Criminal intelligence information service

The *staff perspective* assumed that responsibility and freedom of action had a positive effect on the staff's commitment to their duties and their job satisfaction. Further, it was thought that a high level of commitment and job satisfaction had a positive influence on results. The *employee perspective* was followed up through yearly surveys of the employees and thus focused on long-term changes. The employee questionnaire covered four main areas: responsibility, autonomy, commitment, and job satisfaction. Questions about responsibility aimed to determine whether the staff performed their duties independently, had well defined areas of responsibility, took personal responsibility for their work, and were willing to assume more responsibility. Enquiries about autonomy were pursued to determine whether the staff had enough freedom of action, made decisions without fear, got relevant information, and the extent to which they perceived that superiors trusted them. Under the commitment factor, questions attempted to reveal whether staff deemed their jobs as important, whether they were encouraged to come up with new ideas, could see how their work fit into the overall work of the unit, and got feedback about results, and whether they were able to use their knowledge in their work. Measures of job satisfaction attempted to determine whether the staff were satisfied with their current duties, whether they perceived that their undertakings were correctly assessed by peers and management, whether they found their current duties interesting, whether they co-operated well, and whether they were satisfied at the end of the working day.

The *resource perspective* was intended to represent issues of public funding, management of a unit's finances, personnel development, competence levels of staff, and investments in new technologies.

The *citizen perspective* was intended to capture public opinion about local problems and levels of crime victimization and fear of crime. As noted earlier, annual surveys were conducted to measure police success in "reducing crime and increasing citizens' safety". Accordingly, police needed reliable indicators on citizen safety as well as to learn more about the connection between perceived levels of safety and actual crime levels—as long as these aspects were largely neglected by the NPB's central information system. Therefore, a questionnaire was designed to gather information about these issues, and it comprised four main aspects: problems in residential areas (*e.g.*, littering, criminal damage, drunken people in public places, drug addicts in public places, apartments occupied by drug/alcohol abusers, fighting, women being accosted, teenage gangs, road traffic); crime victimization (*e.g.*, physical violence, theft, criminal damage); fear of falling victim to crime (*e.g.*, theft/criminal damage, assault, fear of going out at night, fear of some individuals or groups, *refraining* from activities in public places); and the extent to which the public considered that the police were doing something to resolve local problems.

Using the Balanced Scorecard

The county level, characterized by chiefs with a limited experience of police work and considerable time constraints that prevented their managing their districts by sight, deemed the development of the scorecard as useful, especially for planning purposes. One police chief noted: "The scorecard has definitely helped me in my daily work, especially in planning activities." Relevance of the scorecard was reckoned higher than reports produced by the central information system. The latter reports were considered

too "cold" and excessively focused on variation analyses that, ultimately, were of little help in explaining the reasons behind reported variations. The scorecard, by contrast, helped them to ascertain the 'whys' of the situation: "The strength of the scorecard comes from its balanced nature. This makes it possible to assess what citizens and staff think with respect to available resources."

Utilization of the balanced scorecard at the county level presented its problems, especially in the aspects of aggregation and comparison. On aggregation, a police chief noted: "The more you aggregate, the more it turns out yellow, and that's a problem for us." In a similar vein, chiefs found problems assessing the performance of individual districts: "If a district reports red on two scorecards and green on one and yellow on the others, what is then the overall picture of the district?" Many pointed out that data from the scorecard should be complemented with information produced by the central information system (such as crime statistics) and accounting reports (e.g., compliance with budgeted expenditures), and this would produce a fair picture of a district's performance.

Local level police also identified some problems with the scorecard. Perceived problems mainly dealt with the extent to which actual performance indicators were measurable as well as with difficulties in cascading primary goals to the operational level. Our interviews with community police officers revealed that they found a considerable gap between what was measured and what actually constituted real police work. For example, an interviewee pointed out: "As a community police officer, you have to know as many people as possible, the more the better. You must listen, get tips, help people, know people, and be around. You can't plan that in advance and measure it on a scorecard. I must be out there working. While others deal with statistics and history, I take care of reality." On a similar theme, another officer wondered: "How can I measure our long-term focused work as well as problem-based police work? When are you supposed to get out of the car and talk to young people who have valuable information? Being able to do that is largely a matter of trust between people and us, and how can the scorecard measure trust building?"

Public order police officers concurred with measurement problems identified by their community police colleagues but also raised issues of goal congruence for local officers. Addressing that issue, an officer reported the following incident: "We got a tip about some drug dealers which involved several days of planning and arrangements, as reducing drug dealing rated high in our scorecards. Alleged criminal offenders were travelling by train from Stockholm. Then, when we were waiting at the railway station, the radio called us out on a priority 1 job, a car accident. We had to abandon the stake-out and leave the station. It was frustrating for us to realize that the dealers would not be arrested. To put it briefly, policing is largely dominated by the need to provide rapid responses."

Crime investigators work daytime Monday through Friday. They normally have experience in community and public order policing. An investigated case must be carefully prepared before they call in suspects and witnesses for interrogation. Since many cases are unique, it is difficult to predict how much time a single case should take. New facts and crimes can come up during the investigation and delay the case until further prosecution. Another cause of delays is the stand-by approach requiring all investigators to respond to priority 1 crimes. They must interrupt what they are doing to

join task forces for acute incidents. Such situations often happen and postpone the balances. Further, investigators must take care of suspects arrested during the night. Though police know these people and how to handle prosecution, the suspects are skilled at denying any knowledge of crimes and so it takes time to handle even a single case. In these circumstances, police officers complained: "It is not easy to find success factors that are really related to our work. In turn, that means that one chooses success factors that are easy to account for. It is easy to count the number of reports in the balances, but that does not depict our achievements. That is only a means of reaching something else."

General discussion

Investigation of organizational performance has overwhelmingly relied on evidence gathered from private sector firms. Nevertheless, the past several years have witnessed increasing interest in enhancing effectiveness and efficiency in the public sector, in turn generating considerable investment in the deployment of performance metrics in such settings. Though extant evidence provides many perceptive insights into the specifics of performance frameworks in public sector organizations, little is known about the measurement of organizational performance in police work. Our investigation drew upon the deployment of the balanced scorecard in police work in Sweden. Our findings may, therefore, contribute to recent research on the design of the scorecard. Further, we identify some problematic issues in the implementation of the scorecard in police work which, taken together, depart from the idealized imagery that characterizes studies addressing the development of performance systems in "pioneering" firms (Vaivio, 1999: 410).

The balanced scorecard used in Swedish Law Enforcement comprised four perspectives: success, citizen, staff, and resources.   This finding suggests some additional considerations. First, Swedish Law Enforcement deployed a scorecard that largely relied on the stakeholder perspective on organizational performance (Atkinson *et al.*, 1997a). As noted by Ittner and Larcker (1998: 233), government agencies depend primarily on support of external constituents and to a lesser extent on actual performance. The main constituents of police groups, we argue, are public opinion, the government, and staff. Whereas the former constituents may be considered "environmental stakeholders", that is, those which define the organizational environment, the latter correspond to "process stakeholders", which should pursue the organizational goals within the contexts defined by environmental stakeholders (Atkinson *et al.*, 1997a). We argue that the citizen and the resources dimensions attempted to account for environmental stakeholders such as the public and the government, whereas the staff dimension reflected the needs and demands of process stakeholders, the police personnel. Accordingly, the four perspectives enshrined in the scorecard used by Swedish Law Enforcement "balanced" measures of internal success as well as internal performance (Bourne *et al.*, 2000: 756). Further, the perspectives comprised in the balanced scorecard targeted the main stakeholders of Swedish Law Enforcement, which ultimately invested police work with support and legitimation. The resulting stakeholder perspective, in turn, resembled developments in performance metrics of both non-profit (*e.g.*, health care centres; see Atkinson and Epstein, 2000a, b) and profit-seeking firms (*e.g.*, chemical industry, banking; see Atkinson and Epstein, 2000a: 24). Lastly, in spite of considerations of the success perspective as the "crucial" one for policing, the development of the

stakeholders perspective at Swedish Law Enforcement did not assign different weights to each of the perspectives as was done by the Bank of Montreal (Epstein and Birchard, 1999).

Second, the score cards drew on annual surveys of public opinion and police personnel to gather data on indicators of the citizen and staff perspectives. Measurement of performance indicators through surveys conformed to the assertion by Bititci *et al.* (2000a: 697) that performance measurement frameworks should use indicators that monitor changes in critical parameters of the external environment. The scorecard used by Swedish Law Enforcement, we argue, tracked its external environment through surveys that monitored citizens' demands with respect to policing and public perceptions about the role of police in enhancing the quality of urban life. Further, under the resource perspective, police staff were annually surveyed about issues of autonomy, commitment, responsibility, and job satisfaction. The outcomes of surveys administered by Swedish Law Enforcement to its main stakeholders, in short, conform to similar developments of local governments (Kaplan and Norton, 2001) and health care centres (Atkinson and Epstein, 2000b).

Third, the resource perspective targeted the needs and demands of the government, a major constituent of Swedish Law Enforcement. The deployment of this perspective in the scorecard resulted in the development of performance metrics that eventually aimed at enhancing resource utilization. Consideration of the resource perspective in the scorecard of the Swedish police may respond to the wave of managerialism that has spread across the public sector and that emphasizes a reconception of performance in terms of "efficiency", "economy", and "effectiveness" (Ballantine *et al.,* 1998: 74; Brignall and Modell, 2000: 289).

Fourth, measurement of organizational performance in the Swedish police organization addresses a triple time perspective (past, present, and future). Bourne *et al.* (2000) point out that a framework of organizational performance should provide an early indication of future business performance as well as a record of what has been achieved in the past. This combination of past and future perspectives, in turn, constitutes a salient characteristic of the balanced scorecard. As noted by Norreklit (2000), the scorecard complements the historical nature of financial metrics with future indicators of performance. In the case of the scorecard used by Swedish Law Enforcement, the future perspective is developed through surveys to both citizens and staff. Such surveys, we contend, provide the police authorities with relevant data about their demands, concerns, and perceptions about the quality of police work. Furthermore, whereas the past perspective is tracked through the monthly release of indicators of success, the present perspective is measured by metrics of available resources.

Lastly, the stakeholder approach to balanced scorecard at Swedish Law Enforcement constituted an antecedent of the "second generation" of performance measurement systems. Compared to the Performance Prism (Neely *et al.,* 2002), the deployment of a performance measurement system to policing in Sweden may be regarded as a preliminary attempt. For Neely *et al.*, the Performance Prism consisted of five interrelated perspectives on performance: stakeholder satisfaction, stakeholder contribution, strategies, processes, and capabilities. In a less systematic way, Swedish Law Enforcement measured the satisfaction of two main stakeholders (e.g., personnel and citizens), as well as resources provided by a third one (e.g., the government). On the

other hand, the success perspective constituted a mixture of strategies and process. The confusing nature of this mixture and an unclear definition of necessary capabilities to cope with the new paradigm of policing could be some of the areas to consider in eventual revisions of the balanced scorecard of Swedish Police Enforcement.

Our results indicate that the implementation of the balanced scorecard at Swedish Law Enforcement experienced problems with measurement of performance metrics and aggregation of data. This finding, we argue, has some additional implications. As noted by Atkinson and Epstein (2000a: 27), performance indicators should conform to criteria of measurability and completeness that characterize organizational performance frameworks. The lasting influence of managerialism on "new" public management, however, results in a celebration of measurability on the part of government agencies which may instil a bias in performance towards the development of easily measurable concepts (Brignall and Modell, 2000). Admittedly, though, what cannot be measured is still relevant (Mooraj, Oyon and Hostettler, 1999). When the scorecard was implemented at Swedish Law Enforcement, some crucial activities of community policing were not measured, such as building relationships with people in the community and building trust with youth. These activities are among those that exemplify the changing paradigm in police work, which shifts emphasis from making arrests to enhancing the quality of urban life (Sparrow, Kennedy and Moore, 1992), and that constitute a salient characteristic of the governmental vision of Swedish Law Enforcement. In spite of the importance attributed to community policing and crime prevention in the Swedish police setting, the scorecard neglected such activities. In contrast, regular activities of the traditional conception of policing were extensively monitored under the success perspective, which in turn was regarded as the "most crucial strategic dimension within the balanced scorecard model [of Swedish Law Enforcement]". Therefore, in spite of the intended modern model of the Swedish police, the scorecard was formed by a set of performance indicators biased towards issues of economy and efficiency (Palmer, 1993), which in turn supported the conventional vision of police work. As Morgan (1986) aptly noted, implementation of performance indicators makes employees adapt their behaviour to the specifics of the metrics in use.

Measurement of performance indicators in Swedish Law Enforcement used a traffic light system: red, yellow, and green, as endorsed by Bititci *et al.* (2000a: 700): "Colours are commonly used in performance measurement software." Aggregation of non-financial measures, though, is regarded as a source of problems in the implementation of balanced scorecards (Bond, 1999). Though traffic light systems are widely accepted as measures of organizational performance, such measurement procedures may well be regarded as coarse-grained. Measurement of policing through colours, we argue, poses insurmountable obstacles to capturing the richness and variety of police work. Therefore, as aggregation progresses, most dimensions turn into yellow, and "yellowness" diminishes the potential of the scorecard for purposes of operational improvements at the local level.

Aggregation of measures, hence, constitutes another potential area for improvement in the balanced scorecard of Swedish Law Enforcement. From our standpoint, however, the use of colours is useful at the initial stages of implementing a performance measurement system in an organization that lacked tradition in the deployment of such measurements. However, as the project makes progress, the use of colours may be replaced by refined, quantitative metrics. Admittedly, the aggregation of data is

inevitably problematic in any performance measurement system. However, recent research provides considerable advances in this area. In particular, Bititci, Suwignjo and Carrie (2000b) use the Quantitative Model for Performance Measurement System to quantify the effects of factors on performance through the Analytic Hierarchy Process (Saaty, 1980). The approach of Bititci *et al.* (2000b) towards the aggregation of performance factors consisted of (i) the determination of cognitive maps that enabled the identification of factors affecting performance and their relationships, (ii) the determination of cause-and-effect diagrams to structure the factors hierarchically, and (iii) quantification of the effects of the factors on performance.

Taking these results together, we found that the implementation of the balanced scorecard in Swedish Law Enforcement provided many insights into the overall process of deploying performance metrics in public sector organizations. The Swedish police adopted a stakeholder approach to meet the demands of major police constituents. Further, the framework itself proved helpful for central officers for planning purposes. However, as noted by Brignall and Modell (2000: 286), implementation of performance measurement systems in public sector organizations, characterized by multiple stakeholders with complex, heterogeneous, intangible services, delivered in circumstances of high uncertainty about means-ends, is especially difficult. In the particular case of Swedish policing, we found that operational, local level officers perceived the system as non-relevant and time consuming. This perception, which largely contrasted with the positive discernment of central officers about the usefulness of the scorecard, may be attributed to lack of measures of some crucial tasks of policing as well as to a questionable relevance of aggregated indicators for police work at the local level.

Extensions

This investigation has some limitations that may encourage future work. First, our study addressed the deployment of the balanced scorecards in Swedish Law Enforcement. The Swedish government enacted community policing and crime prevention as priorities of police work. Surprisingly, however, the Swedish police scorecard raised some caveats in assessing performance of these crucial activities of policing. Investigation of performance frameworks in police organizations that hold a traditional role of policing (e.g., making arrests) may provide a complementary perspective on the role of performance measurement systems in police work assessment. In a similar vein, investigation of balanced scorecards in police organizations that approach the new paradigm of policing (e.g., community policing, crime prevention) may shed some light about the sort of indicators that enable measurement of such complex activities.

Second, our study focused on the implementation of the balanced scorecard in Swedish Law Enforcement, a setting characterized by a *Scandinavian* approach to employees and social issues. As far as the implementation of the scorecard was concerned, though, emphasis on employees was limited to issues addressed in the annual survey of the staff perspective. Strikingly, however, what was not incorporated into the scorecard was any measure that facilitated police work at the lower levels of the chain of command. Accordingly, we deem that future research addressing performance measurement systems in centralized and decentralized police organizations may enhance understanding about the role of the balanced scorecard in rendering effective, efficient and "modern" public sector organizations.

References

Atkinson, A. & Epstein, M. 2000a, Measure for Measure: Realizing the Power of the Balanced Scorecard, <u>CMA Management</u>, September, pp. 23-28.

Atkinson, A. & Epstein, M. 2000b, <u>Applying the Balanced Scorecard</u> (Mississauga, ON: CMA).

Atkinson, A.A., Balakrishnan, R, Booth, P., Cote, J.M., Groot, T., Malmi, T., Roberts, H., Uliana, E. & Wu, A. 1997a. New directions in management accounting research, <u>Journal of Management Accounting Research</u>, Vol. 9, pp. 79-108.

Atkinson, A.A., Waterhouse, J.H. & Wells, R.B. 1997b, A Stakeholder Approach to Strategic Performance Measurement, <u>Sloan Management Review</u>, Spring, pp. 25-37.

Ballantine, J.A., Brignall, T.J. & Modell, S. 1998, Performance Measurement and Management in Public Health Services: A Comparison of UK and Swedish Practice, <u>Management Accounting Research</u>, Vol.9, No 1, pp. 71-94.

Berg, B.L. 1998 <u>Research methods for the social sciences</u> (Boston: Allyn and Bacon).

Bititci, U.S., Turner, T. & Bergerman, C. 2000a, Dynamics of Performance Measurement Systems, <u>International Journal of Operations and Production Management,</u> Vol. 20, No 6, pp. 692-704.

Bititci, U. S., Suwignjo, P. & Carrie, A. S. 2000b, Strategy management through quantitative modelling of performance measurement systems, <u>International Journal of Production Economics,</u> Vol. 69, No. 1, pp. 15-22.

Bond, T.C. 1999, The Role of Performance Measurement in Continuous Improvement, <u>International Journal of Operations and Production Management</u>, Vol. 19, No 12, pp. 1318-34.

Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. 2000, Designing, implementing and updating performance measurement systems, <u>International Journal of Operations and Production Management,</u> Vol. 20, No 7, pp. 754-71.

Brignall, S. & Modell, S. 2000, An Institutional Perspective on Performance Measurement and Management in the 'New Public Sector' , <u>Management Accounting Research,</u> Vol. 11, pp. 281-306.

Brunsson, N. 1994, Politicization and 'Company-ization' – On Institutional Affiliation and Confusion in the Organizational World, <u>Management Accounting Research,</u> Vol.5, pp. 323-35.

Dervitsiotis, K.N. 1997, An Approach Relating Total Performance Improvement with Financial Results, <u>Total Quality Management</u>, Vol.8, No.1, pp. 67-82.

Dilulio, J.J, Wilson, J.Q., Logan, C.H., Cole, G.F., Moore, M.H., Alpert, G.P. & Petersilia, J. 1994, <u>Performance Measures for the Criminal Justice System: Discussion Papers from the BJS-Princeton Project,</u> (Princeton: Diane Publishing).

Epstein, M.J. & Birchard, B. 1999, Counting What Counts: Turning Corporate Accountability to Competitive Advantage (Reading, MA: Perseus Book).

Fitzgerald, L., Johnston, R., Brignall, S., Silvestro, R., & Voss, C. 1991, Performance Measurement in Service Businesses (CIMA: London).

Hood, C. 1995, The 'New Public Management' in the 1980s: Variations on a Theme, Accounting, Organizations and Society, Vol. 20, pp. 93-109.

Ittner, C.D. & Larcker, D.F. 1998, Innovations in Performance Measurement: Trends and Research Implications, Journal of Management Accounting Research, Vol. 10, pp. 205-38.

Kald, M. & Nilsson, F. 2000, Performance Measurement at Nordic Companies, European Management Journal, Vol. 14, No.1, pp. 113-27.

Kaplan, R.S. 1983, Measuring Manufacturing Performance: A New Challenge for Accounting Research, The Accounting Review, Vol. LVIII, No.4, pp. 686-705.

Kaplan, R.S. & Norton, D.P. 2001, Transforming the Balanced Scorecard from Performance Measurement to Strategic Management: Part 1, Accounting Horizons, Vol. 15, No. 1, pp. 87-104.

Kaplan, R.S. & Norton, D.P. 1996, Translating Strategy into Action: The Balanced Scorecard, Harvard Business School Press, Boston.

Kaplan, R.S. & Norton, D.P. 1992, The Balanced Scorecard-Measures that Drive Performance, Harvard Business Review, January-February, pp. 71-79.

Lipe, M.G. & Salterio, S.E. 2000, The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures, The Accounting Review, Vol. 75, No .3, pp. 283-98.

Lynch, R.L. & Cross, K.F. 1991, Measure Up! Yardsticks for Continuous Improvement (Basil Blackwell: London).

Marshall, C. & Rossman, G.B. 1999, Designing Qualitative Research (Thousand Oaks, CA: Sage Publications).

Maskell, B.H. 1991, Performance Measurement for World Class Manufacturing (Cambridge, MA: Productivity Press).

Medori, D. & Steeple, D. 2000, A Framework for Auditing and Enhancing Performance Measurement Systems, International Journal of Operations and Production Management, Vol. 20, No 5, pp. 520-533.

Mooraj, S., Oyon, D. & Hostettler, D. 1999, The Balanced Scorecard: A Necessary Good or an Unnecessary Evil?, European Management Journal, Vol. 17, pp. 481-491.

Moore, M.H. & Poethig, M. 1998, The Police as an Agent of Municipal Government: Implications for Measuring Police Effectiveness., Paper presented at the National Institute of Justice's Police Research.

Moore, M.H., Spelman, W., Estrich, S.R. & McGillis, D. 1985, Dangerous Offenders: The Elusive Target of Justice (Cambridge, MA: Harvard University Press).

Morgan, G. 1986, Images of Organizations (Beverly Hills, CA: Sage Publications).

Neely, A.D. 1999, The Performance Measurement Revolution: Why Now and Where Next?, <u>International Journal of Operations and Production Management</u>, Vol. 19, No. 2, pp. 205-08.

Neely, A.D., Richards, A.H., Mills, J.F., Platts, K.W. & Bourne, M.C.S. 1997, Designing Performance Measures; A Structured Approach, <u>International Journal of Operations and Production Management</u>, Vol. 17, No. 11, pp. 1131-52.

Neely, A.D., Mills, J.F., Gregory, M.J., Richards, A.H., Platts, K.W. & Bourne, M.C.S. 1996, <u>Getting the Measure of Your Business</u>, (Findlay: London).

Norreklit, H. 2000, The Balanced on the Balanced Scorecard –A Critical Analysis of some of its Assumptions, <u>Management Accounting Research</u>, Vol. 11, No 1, pp. 65-88.

Olson, O., Guthrie, J. & Humphrey, C. (eds.) 1998, <u>Global Warning! Debating International Developments in New Public Financial Management</u> (Cappelen Akademisk Forlag: Oslo).

Palmer, A.J. 1993, Performance Measurement in Local Government, <u>Public Money and Management,</u> Vol. 13 (4), pp. 31-36.

Rossman, G.B. & Rallis, S.F. 1998, <u>Learning in the Field: An Introduction to Qualitative Research </u> (Thousand Oaks, CA: Sage).

Saaty, T.L. 1980, <u>The Analytic Hierarchy Process</u>, (New York: McGraw Hill).

Silk, S. 1998, Automating the Balanced Scorecard, <u>Management Accounting-US</u>, May, pp. 38-44.

Sparrow, M.K., Kennedy, D.M. & Moore, M.H. 1992, <u>Beyond 911: A New Era for Policing</u> (Harper and Collins: New York).

Swedish Code of Statutes 1959, <u>General Instructions to the Police</u>. SFS 1959: 492.

Vaivio, J. 1999, Exploring a 'Non-Financial' Management Accounting, <u>Management Accounting Research</u>, Vol.10, No 4, pp. 409-437.

Van Peursem, K.A., Pratt, M.J. & Lawrence, S.R. 1995, Health Management Performance: A Review of Measures and Indicators, <u>Accounting, Auditing and Accountability Journal,</u> Vol.8 (5), pp. 34-70.

Walker Information 1998, A North American Study on Stakeholder Measurement, <u>Measurement</u>, Vol. 8 (3), pp. 1-4.

Wilson, S. 1977, The Use of Ethnographic Techniques in Educational Research, <u>Review of Educational Research</u>, Vol. 47, pp. 245-265.

NOTAS

NOTAS

NOTAS