Improving performance in a Swedish police traffic unit: Results of an intervention

Robert D. Pritchard a, Satoris S. Culbertson b,⁎, Kenneth Malm c, Anders Agrell d

a Department of Psychology, University of Central Florida, Orlando, FL 32816-1390, United States
b Department of Psychology, Kansas State University, 421 Bluemont Hall, 1100 Mid-Campus Drive, Manhattan, KS 66506-5302, United States
c ProMES International Sweden AB, 70357 Örebro, Sweden
d Department of Social Science Psychology, University of Örebro Sweden, 701 82 Örebro, Sweden

A B S T R A C T

This article describes the results of a feedback system designed to improve performance for a Swedish traffic police unit and examined whether such a feedback system was beneficial or detrimental to the attitudes of the officers. As in many Western countries, government organizations are being required to demonstrate their effectiveness with quantitative performance measures. An approach called the Productivity Measurement and Enhancement System (ProMES) was used with three groups of Swedish Traffic Police to do this. ProMES is a method for identifying unit’s objectives, developing measures for these objectives, and using this information as feedback. ProMES was developed with these police units and feedback from the system was used over a four-year period. Results indicate that there were substantial increases in performance. There were also decreases in accidents, injuries, and fatalities compared both to baseline and to comparison groups in Sweden. These improvements were made with fewer and fewer police officers each year.

© 2008 Elsevier Ltd. All rights reserved.

Introduction

As in many Western countries, government organizations are being required to become more accountable by using performance measures to evaluate the contributions made by the agency. In addition, agencies are being increasingly asked to monitor and improve their performance as the need to control budgets and pressures to downsize remind decision-makers of the need to do more with less. While this need to be more efficient while maintaining effectiveness is important, however, there are other factors to consider. For example, it is also important that every employee understands the overall organizational mission as well as the plan to accomplish that mission using departmental goals and objectives (Harrison, 1996). It is not always clear to departments, however, how to accomplish this without destroying morale.

The primary purpose of this study was to evaluate a specific intervention that accomplished these objectives with a sample of Swedish traffic police units. This study sought to examine the results of a measurement and feedback system for improving performance while also increasing alignment of efforts with organizational objectives. In this article, performance is defined as how well the outputs produced by the unit meet organizational objectives. The effects of this intervention on attitudes of the officers were also examined. A secondary purpose of this study was to explore two problematic measurement issues common to many types of work, but especially to police work: control over outcomes and detection of negative events. This article first describes the theoretical background underlying the intervention, followed by a summary of how the intervention is done and a discussion of specific measurement issues. The article concludes with a presentation of the results of the intervention and a discussion of the findings.

Theoretical background

The intervention used in this study was the Productivity Measurement and Enhancement System, or ProMES (Pritchard, 1990, 1995; Pritchard, Harrell, DiazGranados, & Guzman, 2008; Pritchard, Holling, Lammers, & Clark, 2002). ProMES is a results-oriented measurement and feedback system specifically designed to improve performance over time, while at the same time improving the quality of work life. The theoretical background of ProMES comes primarily from the motivational aspects of the Naylor, Pritchard, and Ilgen (1980) (NPI) theory and a more recent motivation theory (Pritchard & Ashwood, 2008) based on NPI theory. These theories are expectancy theories, proposing that individuals are motivated by the anticipation of how their effort will lead to the satisfaction of their needs (e.g., Campbell & Pritchard, 1976; Heckhausen, 1991; Kanfer, 1990, 1992; Latham & Pinder, 2005; Mitchell & Daniels, 2003; Vroom, 1964).

The Pritchard and Ashwood (2008) theory posits that individuals have a certain amount of energy at a given time—their energy pool—that is used to satisfy their needs for such things as food, water, achievement, safety, and power. This energy pool varies across people and across time for any individual. The concept is similar to that of attention resources (Kanfer & Ackerman, 1989; Kanfer, Ackerman,
Murtha, Dugdale, & Nelson, 1994) in that the energy pool concerns the issue of the limited resources individuals have to devote to tasks. Within this framework, motivation is described as the process that determines how this energy is used to satisfy needs.

According to the Pritchard and Ashwood (2008) theory, the motivation process can be broken down into several key components—actions, results, evaluations, outcomes, and need satisfaction. Energy is allocated across possible actions (e.g., a police officer patrolling a neighborhood, issuing citations, or writing reports), which generally produces results (e.g., typing, an action, generates a report, a result). Thus, a result is the individual’s output. When results are observed and an evaluator places the measured result on a good-to-bad continuum, evaluations are produced. Multiple evaluators evaluate the output (e.g., the officer’s report may be evaluated by the officer, his or her supervisor, and/or the district attorney). After these evaluations are made, outcomes occur. These are intrinsic outcomes such as a feeling of accomplishment, or extrinsic outcomes such as forms of recognition, promotion, or pay raises. Outcomes get their motivating power because of their ties to need satisfaction. The more an individual’s needs are satisfied, the greater the positive affect that is experienced.

As with other expectancy theories, the linkages between the components are critical. Within the Pritchard and Ashwood (2008) theory, these linkages are called connections. The first linkage is the actions-to-results connection, which describes the individual’s perceived relationship between the amount of effort devoted to an action and the amount of the result that is expected to be produced. This perceived relationship can range from very strong to nonexistent. The next linkage is the results-to-evaluations connection. This connection reflects the individual’s perceived relationship between the amount of a result that is produced and the level of the evaluation that is expected to occur. There would be such a connection for each different result and for each individual who evaluates the result(s) such as the officer, colleagues, supervisor, prosecutor, and so forth. The strength of these connections varies. The timeliness of reports (a result), for example, may be more strongly related to the supervisor’s evaluation of the officer than the amount of community service. The evaluations-to-outcomes connection is the perceived relationship between the level of the evaluation and the level of outcome expected. The outcomes-to-need satisfaction connection defines the perceived relationship between how much of an outcome is received and the degree of anticipated need satisfaction that will result.

According to Pritchard and Ashwood (2008), the result of these motivation components is the intent to behave. This intent leads to actual behavior, or the application of energy to actions, which in turn leads to actual results, evaluations, outcomes, and need satisfaction. These actual events have a feedback relationship with the various motivational components. For example, actual outcomes received influence subsequent evaluations-to-outcomes connections.

In addition to Naylor et al. (1980) theory and the Pritchard and Ashwood (2008) motivational theory, the development of ProMES has been influenced by several other bodies of literature. These include the literature on feedback, goal setting, participation, roles and role ambiguity and conflict, and team effectiveness. How these literatures influenced the design of ProMES is described below, after the description of ProMES.

The Productivity Measurement and Enhancement System (ProMES)

The next section describes the specific intervention used. This is followed by a discussion of how the intervention operationalizes the theory and other literatures.

Development of the system

The ProMES intervention is typically done in a series of steps, described in greatest detail in Pritchard (1990). To summarize, a design team is formed composed of people from the target unit, one or two supervisors, and a facilitator familiar with ProMES. This design team meets to identify the objectives of the unit and corresponding quantitative measures (indicators) that assess how well the unit is meeting the objectives. The objectives and indicators are then approved by higher management in a formal meeting between the design team and higher management where management reviews, and if necessary, works with the design team to revise the objectives and indicators.

Objectives and indicators

The objectives and indicators might look like the following. (In most actual cases, four to six objectives and eight to fifteen indicators are developed, but to keep the example manageable, only a subset was used.)

Objective 1: respond to emergency calls
Indicator 1: average number of minutes to respond to emergency calls

Objective 2: investigate crime
Indicator 2: percentage of violent crimes leading to arrest
Indicator 3: percentage of violent crimes handled within thirty days

Objective 3: aid in the prosecution of crime
Indicator 4: percentage of arrests transferred to the prosecutor

Objective 4: facilitate crime prevention programs
Indicator 5: number of ongoing prevention programs
Indicator 6: percentage of officer time for educational crime prevention programs

Contingencies

Once the objectives and indicators are approved, the design team develops what are known as contingencies. Contingencies are a type of graphic utility function relating variation in the amount of the indicator to variation in unit effectiveness. In other words, it is a function that defines how much of an indicator is how good for the organization. Fig. 1 shows examples of contingencies for four of the indicators above. The upper left quadrant of Fig. 1 is the contingency for the first indicator, average number of minutes to respond to emergency calls. Varying amounts of this indicator are shown on the horizontal axis, ranging from a slower response time of fifteen minutes to a quicker response time of three minutes. The vertical axis is the effectiveness score. Effectiveness is defined as the amount of contribution being made to the organization. It ranges from -100, through 0 to +100. The zero point is defined as the amount of the indicator just meeting minimum expectations. Indicator amounts above this expected level get a positive effectiveness score. The higher the unit is above this expected level, the higher the effectiveness score. Indicator amounts below the expected level receive a negative effectiveness score.

The contingency relates indicator amounts to the effectiveness scores. For example, the contingency for the first indicator (average number of minutes to respond to emergency calls) shows that the minimum expected level is nine minutes. It is not expected that average response time can be lower than three minutes because of time to transmit the call and the distance to be traveled. The contingency indicates that responding nine minutes and more gets progressively worse. Going from an average of nine minutes to five minutes produces a substantial increase in effectiveness, but responding in less than five minutes produces a much smaller improvement in effectiveness. This is because, although responding faster than five minutes is of value, it could also produce some negative consequences such as driving recklessly to the site. Unit members create a separate...
contingency for each indicator, so in the current example with its four indicators, there are four contingencies.

A formal step-by-step process is followed to develop the contingencies. This procedure is described in Pritchard (1990) and in Pritchard et al. (2008). It essentially consists of group discussion to consensus where contingency development is broken into discrete steps executed by the design team. Each of the different parts of the contingencies are agreed upon, and then put together into a whole. Once the design team has come to agreement on the contingencies, they are presented to higher management for review and approval. This is similar to the step done for objectives and indicators.

Importance of contingencies

Three things are particularly noteworthy about the contingencies. First, they essentially scale the level of output (the indicator level) to how good that is (the effectiveness score). In doing this, they formally define what is considered good, adequate, and poor performance on each indicator. This allows the feedback system to provide both descriptive feedback from the indicator level and evaluative feedback from the effectiveness score. With the contingencies agreed upon, the individuals in the unit and their management know in advance how good or bad each level of output is considered. If a unit gets an effectiveness score above zero, the unit has exceeded minimum expectations. The higher the score, the more they have exceeded expectations. Negative effectiveness scores mean the unit is performing below expectations.

A second feature of the contingencies is that they capture differential importance. Not every indicator is equally important, and the overall slope or range of the effectiveness scores captures this differential importance. For example, Fig. 1 shows that percent violent crimes handled within thirty days is the most important indicator because it has the steepest slope; it ranges from an effectiveness score of -100 to +100. Number of ongoing crime prevention programs is the least important indicator with a range from -30 to +30.

The third noteworthy feature of the contingencies is that they capture nonlinearity. The relationship between how much an organizational unit does on an indicator and the amount of contribution (effectiveness) that level of the indicator makes to the overall functioning of the organization is frequently not linear (Campbell, 1977; Campbell & Campbell, 1988; Kahn, 1977; Pritchard, Jones, Roth, Stuebing, & Ekeberg, 1989; Pritchard, Youngcourt, Philo, etc.).
McMonagle, & David, 2007). It is common, for example, that once the unit’s level of quality reaches a point that satisfies the customer, further improvements in quality are not especially valuable. That is, a point of diminishing returns is reached. The contingencies in ProMES capture this nonlinearity. For example, the contingency in Fig. 1 for percentage of violent crimes handled within thirty days shows a point of diminishing returns after 80 percent. Such a contingency might reflect the belief that because some crimes are not ever going to be fully resolved, devoting the resources to go above 80 percent is not an effective use of resources. The contingency for percent of officer time for educational crime prevention programs (in the lower right-hand quadrant) is a special type of nonlinearity showing that if the percent of officer time goes above 5 percent, the value to the organization actually decreases. It is also important to note that these nonlinearities are very common, with the vast majority of the contingencies developed in ProMES having some degree of nonlinearity.

Feedback using the system

Upon approval of the contingencies by upper management, the feedback system is finished and ready for implementation. Unit personnel collect data on the indicators, and along with the contingency information, feedback is provided to each member of the unit after each performance period during regular feedback meetings. An example of the basic information in the feedback report is shown in Table 1. The indicators are listed along with the score on that indicator for the unit for that period. For example, the unit responded to emergencies in an average of 5.0 minutes for this period. Based on the contingency, this corresponds to an effectiveness score of +20, above minimum expectations. The contingencies rescale all the indicators to a common metric of effectiveness. Thus, they can be added together to produce an overall effectiveness score for that period. In the example, the unit’s overall effectiveness score for this period is +43, well above minimum expectations. Other aspects of the feedback included plots of the overall effectiveness score over time and changes in indicator scores from the past month to the current month.

The feedback meetings are held with all unit personnel and their supervisor to review the feedback report, identify ways to make improvements, and evaluate previous improvement attempts. Whereby the unit personnel essentially design the system, and management approves it, an understanding and alignment of organizational goals and objectives is more likely because any misunderstandings or misalignments would be discussed and resolved. This process of regular feedback reports and meetings goes on over time in a continuous improvement model. The feedback system is reviewed periodically to determine whether changes are necessary.

How ProMES operationalizes the theory

ProMES was designed to operationalize the Naylor et al. (1980) and Pritchard and Ashwood (2008) motivation theories and capitalize on their implications. Results in the theory are the ProMES indicators and results-to-evaluations connections are the contingencies. The actions-to-results connections can be thought of as defining work strategies in that they identify how effort should be allocated across actions. The feedback reports and feedback meetings focus on developing better work strategies, i.e., a more optimal set of action-to-results connections. The feedback over time allows unit personnel to evaluate how well the new strategies are working and to refine them as needed, further refining the actions-to-results connections. The process of developing these connections and having them approved by management ensures role clarity and reduces role conflict because doing the contingencies means that different evaluators agree on what is valued. As employees experience improvements in performance, outcomes are added in the form of recognition from supervisors and feelings of personal accomplishment.

ProMES was also designed to operationalize factors from other literatures that show performance and attitude improvements. The literature on feedback (e.g., Bobko & Colella, 1994; Ilgen, Fisher, & Taylor, 1979; London, 2003; Murphy & Cleveland, 1995; Smith, London, & Reilly, 2005; Taylor, Fisher, & Ilgen, 1984; Wall & Lischeron, 1977; West & Anderson, 1996) has suggested a number of features that an ideal feedback system should include. One such feature is including both descriptions and evaluations of performance, which ProMES does through indicator (description) and effectiveness (evaluation) scores. Another important feature is making the system transparent so people know what their evaluations will be. ProMES accomplishes this through formal agreement of the performance measurement system by all individuals involved in making ProMES the evaluation system. Validity of the measurement system and perceived validity of the system, additional features described in the feedback literature, are maximized through a careful review of the indicators and contingencies by the design team, unit members not on the design team, and upper management. Finally, accuracy of the feedback is ensured through efforts to ensure validity, maximize participation, make the system transparent, and give regular feedback.

ProMES also includes aspects of goal setting (Latham & Pinder, 2005; Locke & Latham, 2002). While goal setting clearly includes formal, relatively public, agreed-upon levels of output to strive for (formal goal setting), it also includes less formal processes such as private and public intentions to act (Frese & Zapf, 1994; Locke & Latham, 2002). ProMES provides feedback for learning goals, and feedback meetings focus on behaviors necessary to attain those learning goals that will help performance (Brown & Latham, 2002; Latham, Mitchell, & Dossett, 1978).

ProMES also includes features from the participation literature. There is considerable evidence that participation on important issues can have positive effects on employee performance and attitudes (Cawley, Keeping, & Levy, 1998; Dipboye & de Pontbriand, 1981; Locke & Schweiger, 1979). Participation enhances perceptions of procedural justice and voice (Cawley et al., 1998; Lind, Kanfer, & Earley, 1990). ProMES design teams participate heavily in formulating the measurement system and are encouraged to discuss the development process with unit members outside of the design team. In addition, the entire unit participates in feedback meetings.

The literature on roles is also addressed within ProMES. Roles in Naylor et al. (1980) and Pritchard and Ashwood (2008) are defined as the set of results-to-evaluations connections. These connections identify expected outputs, indicate their relative value of different measures, and define how level of output is related to value and to evaluations. Role conflict and ambiguity influence performance and attitude variables (Fisher & Gitelson, 1983; Jackson & Schuler, 1985; Tubrè & Collins, 2000). Role ambiguity is reduced by identifying the results-o-evaluations connections; role conflict is reduced by gaining agreement on connections by unit personnel, supervisor(s), and upper management.

Finally, the teams literature is also relevant. In a study examining thousands of teams, West (2007) found teams performed better and were more satisfied when there were clear objectives, members

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Basic feedback information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicators</td>
<td>Indicator value</td>
</tr>
<tr>
<td>Average minutes to respond to emergencies</td>
<td>5.0</td>
</tr>
<tr>
<td>Percent violent crimes handled within thirty days</td>
<td>47%</td>
</tr>
<tr>
<td>Number of ongoing crime prevention programs</td>
<td>15</td>
</tr>
<tr>
<td>Percent officer time spent on prevention programs</td>
<td>6%</td>
</tr>
<tr>
<td>Overall effectiveness</td>
<td></td>
</tr>
</tbody>
</table>
worked together to achieve objectives, and when they met regularly to review team effectiveness and how it could be improved. ProMES includes all three of these characteristics. Additionally, Salas, Kosarzycki, Tannenbaum, and Carnegie (2004) and Salas, Rosen, Burke, Goodwin, and Fiore (2006) identified numerous factors affect team effectiveness including holding shared mental models, having clear roles and responsibilities, engaging in a prebrief-performance-debrief cycle, cooperating and coordinating, and using multiple criteria. ProMES creates shared mental models of the work through the development of the system. Roles and responsibilities are clarified during the development of the system. Additionally, the ongoing feedback meetings are a type of prebrief-performance-debrief cycle, and cooperation and coordination are encouraged through the feedback meetings. Finally, multiple criteria of performance are included in the multiple indicators.

ProMES and other interventions

There are other interventions that have similarities to ProMES such as total quality management (e.g., Mahoney & Thor, 1994) and the balanced scorecard (Kaplan & Norton, 1993, 1996). ProMES has a number of features, however, that make it different from these approaches. First, ProMES has a clear underlying theory that led to its development, Naylor et al. (1980) and the Pritchard and Ashwood (2008) theories. It is also based on other literatures such as feedback, goal setting, roles, and participation. ProMES is also a bottom-up strategy rather than the top-down approach of the balanced scorecard and relies heavily on the participation of the people doing the work. This high level of participation means helps improve the acceptance of the resulting system. The combination of people doing the work developing the system and their ideas for the measurement system reviewed and approved by higher management helps improve alignment with broader organizational objectives. The most unique structural feature is the use of contingencies in ProMES. These capture differential importance of the measures, define what level of performance is expected, and capture nonlinearities. These nonlinearities are especially important because they can be used to quantify and thus clearly communicate the organizational priorities for improvement strategies. For example, consider the two top contingencies in Fig. 1. Suppose for a given month, response time averaged nine minutes and percent violent crimes handled was 80 percent. Improving from nine to seven minutes response time adds about fifteen effectiveness points while improving violent crimes handled from 80 percent to 90 percent adds only about seven effectiveness points. So even though violent crimes handled is overall the more important measure, it is better to focus improvement efforts on response time. The determination of effectiveness improvement for different measures allows for improvement strategies to be made where they can have the most positive effect on overall effectiveness. Finally, as described in the next section, there is clear evidence that ProMES has a positive effect on performance based on applications by many scholars, in different types of jobs, different types of organizations, and in different countries. There is surprisingly little scientifically acceptable evidence that other interventions actually improve performance.

Past experience with ProMES

Although ProMES has been shown to be effective in many settings (Pritchard et al., 2008), prior to this study it had not been used specifically with law enforcement personnel. Nevertheless, Pritchard et al. (2008) reported results from eighty-three ProMES interventions where the overall effectiveness score during baseline was compared to that during feedback. Results showed large improvements under feedback. The mean effect size from baseline to feedback was 1.16 and 1.44 when weighted by number of time periods in each study. Put another way, performance improved under feedback by 1.16 standard deviations, 1.41 when weighted. These effect sizes are very large, and assuming performance over time forms roughly a normal distribution, mean that overall effectiveness during feedback equals what was the eighty-eighth percentile under baseline. The weighted effect size of 1.44 is the ninety-third percentile under baseline. Performance data for control groups who did not get ProMES feedback showed no change in performance during the time ProMES was done. These findings led to the prediction that ProMES will improve performance in criminal justice settings.

Hypothesis 1. Performance will be higher during ProMES feedback than during the baseline prior to feedback.

Specific measurement issues

Two specific measurement issues must be addressed in any measurement of police performance, whether ProMES or another approach is used. A special focus of this project was to attempt to deal with these issues in a novel way.

Control over outcomes

The ultimate goals of police work are often not under the control of the officers. For example, typical goals include preventing crime, enhancing public safety through education and enforcement, and safeguarding property and individual rights. There are numerous factors, however, that police officers cannot control that contribute to crime rates, such as the socioeconomic status of individuals within their beats (Bukanya, 2005; Rutter, Giller, & Hagell, 1998), pollution levels (Rotton, 1983; Rotton & Frey, 1985), population density of their city (Christens & Speer, 2005), the local job market (Wang, 2005), or even the weather (C. A. Anderson, 1989; Hipp, Bauer, Curran, & Bollen, 2004). Furthermore, as Eisenberger, Fasolo, and Davis-LaMastro (1990) noted, patrol officers typically lack discretionary control over many aspects of their performance such as the number of occasions they are called to the scene of accidents or the time spent in court waiting to testify. This creates a dilemma when assessing police performance. Officers want to be evaluated on factors under their control, but there may be few that truly fall under their control, and the factors under their control may not directly influence the ultimate goals of public safety.

This importance of being evaluated on measures that are controllable is supported by research and theory. A variety of sources have suggested that the less control one has over outputs, the lower his or her motivation is to improve those outputs (Frese & Zapf, 1994; Muckler, 1982; Spector, 1986; Wall, Corbett, Martin, Clegg, & Jackson, 1990). For example, the Pritchard and Ashwood (2008) theory would predict that if the relationship between effort and performance is low (action-to-results connections), motivation will be low. Not having control over the measure means personnel cannot impact the measure by their efforts. Furthermore, the importance of control over outputs goes beyond motivation. Indeed, in a meta-analytic review of the literature on control, Spector (1986) found that high levels of perceived control are positively related to motivation, job satisfaction, commitment, involvement, and performance.

The dilemma, therefore, is whether to measure those things that are the most important to the organization and let motivation, job satisfaction, commitment, work involvement, and performance suffer, or measure those things the people doing the work have control over, but which are not of primary value to the broader organization. This dilemma was present in this project. For the Swedish government and National Board of the Police Force (Rikspolitistyrrelsens), the most important outcomes for traffic police were traffic accidents, injuries, and fatalities. These objectives, however, are not easily under the control of the officers.

To deal with this issue, the design team specified objectives and indicators under their control, but which they believed would
subsequently lead to the ultimate goals of decreased accidents, injuries, and fatalities. To address these ultimate goals, they agreed upon three overall objectives: enforce compliance with speed limits, decrease motorists’ use of alcohol or other drugs, and decrease hazardous driving behavior. These objectives, however, were still not under their control. Therefore they created indicators under their control that would ultimately help them reach their objectives, and subsequently the goals set by the Swedish government and National Police Board. For example, although they had no control over motorists’ actual drug or alcohol usage (Objective 2), they had control over how frequently they monitored specific roads at times when people were most likely to be driving under the influence of drugs or alcohol. This led to the creation of indicators under their control but that they theorized would lead to attainment of the overall objective of decreasing accidents, injuries, and fatalities. The objectives and indicators developed by the traffic unit are shown in Table 2.

Thus, the Swedish traffic officers designed their system using what they perceived to be controllable measures. The plan, however, also called for monitoring the less controllable measures: accidents, injuries, and fatalities. Additionally, some indicators were decidedly not under the control of the officers. For example, as shown in Table 2, another indicator designed to measure Objective 2 (decrease motorists’ use of alcohol or other drugs) that was not within their control included the “number of motorists failing alcohol or drug tests in the last four months.” The logic was the ProMES system could help them make improvements on the things they could control. If they were able to improve on the controllable measures, and the ultimate outcomes of accidents, injuries, and fatalities decreased, this would tell them their “theory” was correct and would meet the goals of the government and the National Police Board. In contrast, if they improved on the controllable measures, but the ultimate outcomes did not improve, it would indicate their theory was not correct and they would then need to change their measurement system. They would have to change their theory of how to have positive impact on the ultimate outcomes. This resolution was accepted by officers and higher management.

### Table 2
Traffic police objectives and indicators

<table>
<thead>
<tr>
<th>Objective 1: enforce compliance with speed limits</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent time checking effectively for speeding out of total amount of time devoted to traffic control</td>
<td></td>
</tr>
<tr>
<td>Percent of speeding violations less than 15 km/h above limit, in relation to total violations</td>
<td></td>
</tr>
<tr>
<td>Percent actual surveillance time to planned surveillance time on defined roads</td>
<td></td>
</tr>
<tr>
<td>Percent surveillance during critical time periods for defined roads</td>
<td></td>
</tr>
<tr>
<td>Percent of vehicles driving within the speed limit</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective 2: decrease motorists’ use of alcohol or other drugs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of total alcohol and drug tests done between 9:00 p.m. on Friday and noon on Sundays</td>
<td></td>
</tr>
<tr>
<td>Percent of total alcohol and drug testing time done at high risk places or on high risk people</td>
<td></td>
</tr>
<tr>
<td>Percent of alcohol and drug tests done when stopping a driver for a traffic offense</td>
<td></td>
</tr>
<tr>
<td>Number of motorists failing alcohol or drug tests in the last four months</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective 3: decrease hazardous driving behavior</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of total traffic monitoring time spent monitoring hazardous passing</td>
<td></td>
</tr>
<tr>
<td>Percent of total traffic monitoring time spent monitoring driving too close to the vehicle in front</td>
<td></td>
</tr>
<tr>
<td>Percent of total traffic monitoring time spent monitoring vehicles crossing roads unsafely</td>
<td></td>
</tr>
<tr>
<td>Number of reported hazardous driving incidents in the last four months</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Objective 4: do efficient police administration</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of correctly and legibly completed traffic violation reports</td>
<td></td>
</tr>
<tr>
<td>Percent of time spent on internal administration in relation to total amount of time available (lower is better)</td>
<td></td>
</tr>
<tr>
<td>Percent of correctly and legibly written reports from vehicle inspections</td>
<td></td>
</tr>
<tr>
<td>Percent of investigations completed within thirty days</td>
<td></td>
</tr>
</tbody>
</table>

#### Hypothesis 2
If overall effectiveness scores improve during ProMES feedback, external measures of accidents, injuries, and fatalities will also improve.

### Detection of negative events

The other measurement challenge addressed in this project deals with measures that detect negative events. Examples are crimes detected by police, near misses detected by air traffic controllers, inappropriate findings detected by accountants, and errors made in clerical work. If the unit being evaluated is responsible for detecting these negative events, this creates a problem because if the unit personnel do a better job of detecting these negative events, the rate of negative events would be expected to increase, at least in the short term. For example, if air traffic controllers do a better job of detecting and reporting near misses, more near misses will likely be reported. This gives the impression the controllers are doing a poorer job because the rate of near misses goes up. A similar issue is present in police work. Examples are frequency of drivers exceeding speed limits, number of motorists driving under the influence of illegal substances, and frequency of robberies. Better detection of these events has the apparent effect of increasing their frequency. This apparent increase in the frequency of negative events can lead to perceptions of poorer officer performance. This creates the dilemma for the officers that doing a better job of detecting negative events may lead to the perception of poorer performance.

This was an issue in this project because several indicators measured negative behaviors: “percent of vehicles driving within the speed limit,” “number of motorists failing alcohol or drug tests in the last four months,” and “number of reported hazardous driving incidents in the last four months.” The officers were concerned about including these measures because the better job they did detecting them, the worse their performance would appear, at least in the short term. The solution they developed was to measure both the frequency of these negative behaviors and directly measure the quantity and quality of their detection efforts. Examples of such detection indicators, shown in Table 2, are “percent of actual surveillance during critical time periods for defined roads,” “number of alcohol and drug tests per month,” and “percent of total alcohol and drug testing time done at high risk places or on high risk people.” Their logic was that by including their detection efforts and the frequency of the negative events, the measurement system would show that increased detection was due to better monitoring. This solution was accepted by everyone involved. There were no specific hypotheses about the negative events issue, but this issue is discussed in detail in the Discussion section.

### Team climate perceptions

Although performance improvement is certainly important, it should not be at the expense of employees. Furthermore, not every individual or department may be open to change. If individuals are not open to change or if they do not benefit from the introduction of a new intervention, then large gains in performance would not be expected. Therefore, whether officer attitudes toward their unit and the department and the climate for innovation were improved or worsened following the implementation of ProMES were examined. In order to do this, the team climate inventory (TCI) developed by West (1990) was used, which measures four dimensions related to group process and the climate for innovation (N. Anderson & West, 1998; West, 1994). Innovation is defined as “the intentional introduction and application within a role, group, or organization of ideas, processes, products or procedures, new to the relevant unit of adoption, designed to significantly benefit role performance, the group, the organization or the wider society” (West & Farr, 1989, p. 16). In this case, the innovations were the changed ways the unit did its work because of ProMES.
Four dimensions of the TCI were expected to be influenced by the process of doing ProMES. The first dimension, vision, refers to how clearly defined, valued, shared, and attainable are the team’s objectives (N. Anderson & West, 1996). By clearly defining and sharing the organizational vision, employees become more empowered to assume authority over their own actions toward the vision (Nanus, 1992), which are directed toward relevant aspects of performance and away from irrelevant or undesired aspects (Kouzes & Posner, 1995; Locke et al., 1991). The process of doing ProMES was hypothesized to be likely to influence vision and shared objectives. In a typical ProMES intervention, unit personnel agree on the measures used to assess whether objectives are being met. Once completed, they are presented to higher-level management for approval. This process therefore was expected to produce a climate where goals are clear, agreed, accepted, and seen as valuable. Thus, it was expected that vision would improve from before the start of the design team’s work to after feedback has been in place.

Participative safety, the second dimension of the TCI, reflects the degree to which group members feel comfortable and secure with their fellow group members to share information and involve each other in decision-making (N. Anderson & West, 1996). It was expected that participative safety increases on a ProMES project because of the highly collaborative nature of the system. Specifically, the majority of the design team members were actual employees in the unit. During this process, the facilitator was responsible for ensuring all team members had an opportunity to express their views. By providing the support to contribute to discussions, the facilitator was likely creating an atmosphere in which team members feel safer in expressing their own views. Furthermore, developing and using the system typically uncovers conflicts and the facilitator helps the group work through these conflicts in a constructive way. After feedback starts, all unit members tend to be heavily involved in the feedback meetings. Thus, there is usually ample opportunity over an extended period of time to learn participative skills and become comfortable with them. This led to the prediction that ProMES would have a positive impact on participative safety.

The third factor is task orientation, which is “the commitment of the team to achieve the highest possible standards of task performance, including the use of constructive progress monitoring procedures” (N. Anderson & West, 1996, p. 59). ProMES was expected to influence this dimension because shared concern for performance would be fostered by identifying and agreeing on the objectives, the indicators, and the contingencies. Furthermore, during the development process, the facilitator is typically responsible for ensuring discussions focus on task relevant issues. This process produces considerable involvement by unit personnel and considerable interest in how well they will perform on the group level measures. Appraisal and constructive challenges to group performance would likely occur in the feedback meetings. Thus task orientation was expected to increase as a result of ProMES.

The final factor predicting team innovation is support for innovation. This dimension includes expectations and approval for and support of attempts at innovation. According to N. Anderson and West (1996), this is the “degree of practical support for innovation attempts contrasted by the rhetoric of professed support by senior management” (p. 59). This dimension was expected to be higher because of the commitment needed by both upper management and the employees actually designing the system. By committing to doing ProMES, the group typically goes through a process that should develop group norms for innovation because the major reason for the discussions in the feedback meetings are to make such changes and innovations. There is support for this not only by the previous experience of working together to develop a measurement and feedback system that everyone accepts, but by having the time, place, and the information to develop and evaluate these innovations. Thus, support for innovation was expected to increase in a ProMES project.

**Hypotheses 3–6.** Employees have higher scores on the vision (Hypothesis 3), participative safety (Hypothesis 4), task orientation (Hypothesis 5), and support for innovation (Hypothesis 6) dimensions during feedback than during baseline.

**Method**

**Participants and procedure**

The participants in this study constituted the entire traffic police personnel in Örebro County, an area near the center of Sweden, due west of Stockholm. The implementation of ProMES started during the autumn 1996, and was done with three teams or shifts of police officers in the traffic division, a total of thirty-three people at the start of the project. The officers in the three separate shifts were responsible for monitoring compliance with speed limits, investigating traffic accidents, checking for motorists’ substance abuse, and generally maintaining safety on the highways in this county.

The ProMES process used with the three traffic units was that summarized in the introduction earlier and followed the procedures outlined in Pritchard (1990). The specific details of the process and the resulting system are described in Agrell and Malm (2002) in more detail. The single design team was composed of representatives from the three teams/shifts and the chief constable, and worked closely with the National Board for Road Maintenance (Vägverket), which is in charge of road safety. This design team met four hours a week for three months to develop a single set of objectives, indicators, and contingencies that would apply to all three groups. The system was then approved by management and the police officers in the units as well as by the Vägverket representatives. The actual feedback was separate for each of the three units. Each reported their own indicator scores, each had their own feedback reports, and each had separate feedback meetings. The four objectives and eighteen indicators developed by these units are shown in Table 2. Feedback was given through a computer program designed to be used with ProMES called ProMES Navigator (more information is available from the Web site http://promes.cos.ucf.edu/computerprogram.php). This program provides both for entry of indicator data and a variety of types of feedback reports.

The three units received their first feedback reports in January 1997. Thus, data collected during 1996 were considered baseline data and data collected in 1997 and after were considered data under the feedback condition. The units used ProMES feedback for approximately five years, until the end of 2001.

**Measures**

**Team climate for innovation**

Team climate for innovation was measured with the Swedish version of the TCI (Agrell & Gustafson, 1994; N. Anderson & West, 1998). TCI questionnaires were given to officers in the three units before the ProMES design team started their work (in 1996), again after one year, at which point ProMES feedback had been operating for six months (in 1997), and finally after several years of feedback (2001).

Each questionnaire contained thirty-eight items related to the four factors. Example items measuring vision included, “How worthwhile do you think these objectives are to the organization?” and “How clear are you about what your team objectives are?” Sample participative safety items included, “People feel understood and accepted by each other,” and “Everyone’s view is listened to, even if it is in a minority.” Sample task orientation items included, “Do members of the team build on each others ideas in order to achieve the best possible outcome?” and “Do you and your colleagues monitor each other so as to maintain a higher standard of work?” Sample support for innovation items included, “People in this team are always searching for fresh, new ways of looking at problems,” and “Team members...
provide practical support for new ideas and their applications.” Respondents were required to indicate the extent to which they agreed with each item on a scale ranging from one (strongly disagree) to five (strongly agree). The Swedish TCI is completed individually but is conceptualized and operationalized as a group-level measure. Thus, the team climate score for each of the four factors was the mean score across the individuals in that group for that factor.

Traffic accidents, injuries, and fatalities
Data on traffic accidents, injuries, and fatalities were collected from Örebro County as well as the rest of Sweden. Data were collected during the baseline period (1995 and 1996) and during the feedback periods from 1997 to 1999. After 1999, however, the data were no longer comparable to earlier years. New methods of collecting and defining these events were put into place. In addition, the construction of new highways substantially changed traffic patterns making comparison with previous data problematic.

Results
Effects of the ProMES intervention
The first hypothesis dealt with whether ProMES feedback would improve performance. Fig. 2 shows these results. The horizontal axis is time showing baseline data available for five months, from August 1996 to December 1996. Feedback started in January 1997 and data were available from then until October 2001.

The vertical axis is the overall effectiveness score averaged across the three units. This is the sum of the effectiveness scores from all the indicators. Fig. 2 indicates that in the months prior to the start of ProMES feedback, overall effectiveness was low. That is, their overall effectiveness scores were negative and well below the minimum expected level; an overall effectiveness score of zero. In fact, the average baseline score is -520.4. Once feedback started, however, there was a rapid rise in overall effectiveness during the first six to eight months of feedback. This increase was maintained with some variation throughout 1998, 1999, and the first months of 2000. The mean overall effectiveness score for the feedback period was 276.7. Each of the three traffic police units showed results similar to those shown in the figure.

Another way to look at the magnitude of the improvements is to look at effect sizes. Effect size used was d (Hunter, Schmidt, & Jackson, 1982). To calculate the effect size, the mean difference in the overall effectiveness score between feedback and baseline was calculated, which was then divided by the pooled standard deviation. These effect sizes were 2.78, 4.19, and 3.22 for the three units. Put another way, the mean overall effectiveness score with ProMES feedback was 2.8 to 4.2 standard deviations higher than it was during baseline. An effect size of .8 is considered large (Cohen, 1992), so the values obtained in this study were very large.

Accidents, injuries, and fatalities
Hypothesis 2 stated that if overall effectiveness scores improve during ProMES feedback, external measures of accidents, injuries, and fatalities also improve. Figs. 3 and 4 show the changes in traffic accidents, injuries, and fatalities over the time periods from 1995 to 1999 for Örebro County and the rest of Sweden. Fig. 3 shows these outcomes of interest in terms of the mean number of accidents, injuries, and fatalities for Örebro County compared to the rest of Sweden. Fig. 4 shows the same information as a percent of baseline.

As shown in Fig. 3, the data indicate that the number of accidents, injuries, and fatalities decreased after feedback started and continued to further decrease each year of the feedback. First, regarding accidents, the fact that the average Swedish county had more accidents is not surprising because Örebro County is smaller than the average Swedish county. What is important, however, is the change in the number of accidents from 1995-96 to 1997-99. Specifically, the number of accidents during the feedback period was 88 percent of what it was during baseline for the average Swedish county but was 82 percent of what it was during baseline for Örebro County. Thus, the decrease in accidents in Örebro County decreased a bit more than did the accidents throughout Sweden.

This can also be viewed as a percentage of baseline, which makes it easier to compare the baseline to feedback and the Örebro County results with the rest of Sweden. To compute this, the mean number of accidents during the two-year baseline (1995-96) was calculated. Then the values for each of the subsequent years were calculated as a percentage of this baseline. This was done separately for all the outcome measures for Örebro County and for the rest of Sweden. As shown in Fig. 4, both the Örebro County units and the rest of Sweden start, by definition, with a value of 100 percent in the baseline years (i.e., the mean of the two years of baseline for both is set at 100 percent). Then, the ProMES units and the rest of Sweden both decrease in accidents and decrease the same percentage the first year of feedback. After that, the ProMES units continue to decrease, whereas the rest of Sweden stays roughly equal to the 1997 values. By 1999, the ProMES units’ accident rate is 72 percent of what it was in baseline, while it was 87 percent for the rest of Sweden. This decrease in traffic injuries for Örebro County was significant (χ² = 34.82, p < .001). Thus, Hypothesis 2 was supported in terms of an improvement in reported accidents.

Similar results emerged for injuries. Specifically, for Örebro County, injuries under the feedback period were 90 percent of what they had been during baseline. For the rest of Sweden, the mean number of injuries increased; during feedback injuries were 102 percent of what they were in the baseline years. Thus, the number of injuries decreased with ProMES feedback and decreased relative to the rest of Sweden. Fig. 4 shows the same results, but viewed as a percentage of baseline. Örebro County had fewer and fewer traffic injuries each year of feedback compared to their baseline and compared to the rest of Sweden. For injuries in 1999, the ProMES units in Örebro County were at 87 percent of their baseline values whereas the traffic units in the rest of Sweden were at 106 percent of their baseline. So whereas injuries in Örebro County decreased over time, the number of injuries in the rest of Sweden increased. This decrease in traffic injuries for Örebro County was significant (χ² = 15.62, p < .001), thus supporting Hypothesis 2 in terms of an improvement in reported injuries.
Finally, similar results emerged for traffic fatalities, although the decreases were not statistically significant ($\chi^2 = 2.38, \text{ns}$), possibly due to the relatively small numbers of fatalities in general. Nevertheless, as shown in Figs. 3 and 4, the number of traffic fatalities in Örebro County decreased with feedback compared to baseline, with the number of fatalities during feedback being 72 percent of what it had been during baseline. For the rest of Sweden, however, the mean for the baseline period was equal to the mean for feedback, with an increase to 118 percent of baseline by 1999. Thus, fatalities for the ProMES units decreased during feedback whereas the fatality rate for the rest of Sweden did not change or got worse. Thus, Hypothesis 2 was not supported in terms of an improvement in reported fatalities in a statistical sense, but it could be argued that there was an improvement in a practical sense.

Number of officers

Although no hypotheses were made related to the number of officers during the intervention the data were instructive. The bottom right quadrants in Figs. 3 and 4 show the number of police officers in Örebro County over the same time period as the outcome measures. As with the rest of Sweden, the number of police officers was decreasing due to budget decreases. In Örebro County, the number of traffic police officers decreased from baseline ($M = 32.5$) to the feedback period ($M = 26.7$). By 1999, the number of officers was down to twenty-three. Although not shown on the figure, the number of police in the rest of Sweden also decreased, although actual numbers were not available. As shown in Fig. 4, by 1999 the number of officers was 71 percent of what it had been during baseline. Thus, Örebro County had a substantial loss of officers, yet substantially improved their performance.

Team climate perceptions

Hypotheses 3–6 concerned the effects of the ProMES intervention on the four TCI dimensions. For each dimension, planned pairwise group comparisons were conducted to compare the team climate perceptions during baseline and feedback. A priori comparisons were made for each of the four dimensions between scores obtained in 1996 (baseline) and the average of those obtained in 1997 and 2001 (feedback). For each of the four dimensions, it was hypothesized that the scores during feedback would be greater than those during baseline. Results yielded support for all dimensions except support for innovation. Specifically, for participative safety, the mean score
increased from 3.63 to 4.10, $t=3.62$, $p<.001$. Vision also increased, from a mean score of 3.67 to 3.98, $t=2.96$, $p<.01$. Task orientation increased from a mean score of 3.57 to 3.81, $t=2.03$, $p<.05$. Support for innovation, although increasing from a mean score of 3.44 was not a significant increase, $t=1.26$, ns. Thus, Hypotheses 3, 4, and 5 were supported whereas Hypothesis 6 was not supported.

In addition to testing for statistical significance, the treatment effect omega-squared ($\omega^2$) was computed to examine whether differences were meaningful between the scores obtained during baseline and feedback periods. Scores of .01, .06, and greater than .15 represent small, medium, and large treatment effects, respectively (Cohen, 1977). The effect size was large for participative safety ($\omega^2 = .15$) medium-to-large for vision ($\omega^2 = .12$), and medium for task orientation ($\omega^2 = .06$). Interestingly, although not statistically significant ($p>.05$), the effect size for support for innovation was medium-to-large ($\omega^2 = .08$).

**Discussion**

The primary purpose of this study was to determine the effects of the ProMES intervention on performance and attitudes. The ProMES effectiveness scores indicated large gains in performance and the measures external to the system (accidents, injuries, and fatalities) decreased relative to baseline and relative to the rest of Sweden, although the reduction in the number of fatalities was not statistically significant. This was all occurring despite the decreases in the number of officers over time.

The results suggest that ProMES is an effective way of responding to the governmental requirements for measuring police effectiveness and also increasing performance in the face of reduced personnel. It also helps align the efforts of the officers with the broader goals of the organization by the way the measurement system is developed. The objectives, indicators, and contingencies are reviewed by higher levels of management and a key issue is how well they are aligned with broader organizational goals. Once the measurement system is approved, the resulting feedback system provides information on how to allocate resources so as to maximize contribution to the organization.

Officer attitudes toward their unit and their department and the climate for innovation showed that the scores during feedback were more positive than during baseline, with the exception of the support for innovation dimension. Thus, the feedback provided by the ProMES
intervention led to increases in vision, participative safety, and task orientation.

The problematic measurement issues of control over output measures and detection of negative events were explored. Control over outcomes was addressed by having the design team develop indicators they did control and that they believed would lead to the ultimate organizational goals of decreasing accidents, injuries, and fatalities. They did improve on their measures and as they predicted, the less controllable measures of accidents, injuries, and fatalities also improved. The officers were satisfied that they had control over their measures and managers in the broader organization were satisfied with the improvements in ultimate objectives.

Although no actual data on the issue of detection of negative events were presented, information was presented regarding how the officers addressed this in their own situation, by including their detection efforts and the frequency of negative events. This solution was accepted by all involved, and may prove to be a useful solution to practitioners facing similar dilemmas.

**Practical considerations**

It is useful to consider why these improvements occurred. While the information available is anecdotal, it is quite consistent. This was the first time these officers took the time to sit down and clarify their vision and what objectives were really important in doing the work. The combination of doing this and actually developing the measures was difficult and time-consuming. By the end of this process, however, the officers had quite a different idea of how to do their work. By pooling their knowledge and experience, they realized that a better strategy would be to focus on those things that would have the greatest impact such as patrolling more frequently on some roads, at certain times, and on certain days. They also saw the necessity for completing paper work accurately and minimizing the time spent on administration.

They were quite surprised when the first feedback indicated that they were not actually doing these things very well. This led to many attempts to change the way they did the work. They were then able to assess how good the new strategies were by studying the subsequent feedback reports. Changes in strategy that led to improvements were kept, changes that did not have an effect were revised. This process led to large improvements in their feedback scores. The steadily increasing feedback scores led to considerable positive affect among the officers. Over time when it became clear that accidents, injuries, and fatalities were going down, this produced further positive affect and increased their desire to perform well.

Another issue to consider in evaluating the results is the time taken to develop the system. The design of the measurement system, getting management approval, and finalizing the feedback system took approximately three months, with the design team meeting four hours a week. This is longer than usual, with the average design team meeting time thirty to thirty-five hours (Pritchard, 1995). There were three different units, however, represented in the design team rather than the usual one unit. Thus, the longer time is not surprising.

Although the time to develop the system was longer than usual, the time it took to enter the data during feedback was minimal, especially once the department implemented the use of software designed to help manage the process and provide reports in a variety of formats. Anecdotally, the feedback was also more meaningful, as the feedback was more readily available and the variety in how the data could be viewed (e.g., overall effectiveness scores, separate indicator scores, shifts combined or separate) allowed greater understanding.

**Limitations and future directions**

Although the team climate perceptions of vision, participative safety, and task orientation improved during ProMES compared to baseline, the reason they increased can only be speculated. It would seem that the ProMES process itself, with its participative aspects and process of clarifying roles and expectations would be an influencing factor in these improvements, but these claims cannot be made without further examination of the intervening variables and thought processes of the officers themselves. Future researchers and practitioners who use ProMES, or any intervention aimed at increasing performance or team climate perceptions (or both), would do well to consider some of the explanatory variables in their design of the system.

The focus of the current intervention was on the overall performance improvement of the group of officers, and not on individuals within the department. Thus, no individual differences were assessed in this study’s measurement. Individual differences, however, could have impacted some of the results found within the current study. For example, individuals vary on motivation, intelligence, and personality traits. Each of these could influence how much effort they put into a task or their perceptions with an outcome. Furthermore, it is not possible to say whether the individuals who left and those who remained (or joined the department) during the intervention were different from one another. Nevertheless, results were separated by time period (baseline versus feedback), and had there been an influencing effect of individual differences, it would be expected that they would counter each other as such variables are frequently normally distributed and likely to appear in both conditions. Thus, although these issues could not be addressed in the current study, this was not perceived to be a substantial limitation to the current study’s design, nor did it detract from the study’s findings. Additionally, although the number of officers was modest, this should not discredit or diminish the findings of this study for several reasons. First, although there were only thirty-three officers at the start of the project, this was all of the traffic police from this county (so not an issue of response rate). Second, there was no reason to believe that the make-up of the officers in this particular county was any different than the make-up of the rest of Sweden. Third, a critical issue was whether the findings would generalize to other officers within Sweden, but also to other officers and departments outside of Sweden. As noted, ProMES has been effective in many settings in many countries. Thus, the positive results found within this study were consistent with a broad pattern of similarly positive results. Therefore, although the specific variables that were examined in this study were different than those examined in other studies, the odds of these findings generalizing to other situations and departments was higher based on the similarly positive results in other studies (e.g., comparably high effect sizes). Finally, information was presented in terms of actual numbers (e.g., accidents, injuries, and fatalities) for both Örebro County as well as the rest of Sweden, but also in terms of percent of baseline. This was done in order to account for the fact that Örebro County is smaller than the average county, as well as to make meaningful comparisons that were independent of sample size.

Finally, despite the improvements in the performance of the officers, the intervention was ultimately discontinued, and thus limited the length of time that could be examined. There were several reasons the project ended, including a new supervisor whose measurement system included a number of new projects (so ProMES did not count), and most importantly, when extra money was made available for salary increases, the units doing ProMES did not get any extra for their improvements in performance with fewer personnel. This caused a negative reaction and loss of enthusiasm for the effort leading to ending it. Based on this, organizations that choose to implement any innovation should consider the reward structure that is set in place to support and encourage effective performance based on the innovation. Clearly, alignment of performance and goals is important; similarly, alignment between performance and reward structures is clearly essential for continued success of an intervention.
Conclusions

In conclusion, this article describes the results of a feedback system designed to improve performance for a Swedish traffic police unit and examines whether such a feedback system is beneficial or detrimental to the attitudes of the officers. Results indicated that there were substantial increases in performance and significant decreases in accidents, injuries, and fatalities compared to both baseline and to comparison groups in Sweden, all while number of officer was decreasing.

Acknowledgements

This article expands on work presented earlier in a chapter by Agrell and Malm (2002). Additional data are presented, more appropriate analyses are used, additional measurement issues are addressed, and implications of the results for criminal justice settings are presented.

Note

1. For further information on the specifics of the program ProMES Navigator, contact Kenneth Malm at Kenneth.Malm@promes.se.

References


