

CPB Document

No 31

May 2003

Performance contracts for police forces

Ben Vollaard

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 90-5833-125-3

Contents

Preface	7
Executive summary	9
Samenvatting (summary in Dutch)	13
1 Introduction	17
1.1 Rationale	17
1.2 Research question	18
1.3 Contribution to policy making	18
1.4 The Interior Department as principal	19
1.5 Approach	20
1.6 Structure	21
1.7 Defining the object of analysis	22
1.8 Scope	22
2 The government's objectives for regional police work	27
2.1 Aim and structure	27
2.2 Translating tasks into objectives	27
2.3 Reasons for the regional police forces' level of discretion	28
2.4 The dimensions of overall police performance	30
2.5 Conclusions	32
3 Realizing the objectives	33
3.1 Aim and structure	33
3.2 The Department's challenge	33
3.3 Incentives from the Department's policies	35
3.4 Other incentives	41
3.5 Conclusions	43
4 Empirical evidence on police performance	45
4.1 Aim and structure	45
4.2 Hypotheses	45
4.3 Crime clear-up rate Nordrhein-Westfalen vs. the Netherlands	45
4.4 Weighted solved crimes per officer across forces	46

4.5	Overall productivity of the police, 1980-2000	48
4.6	Possible explanations for the decline in efficiency	50
4.7	Conclusions	55
5	The rationale for performance contracts	53
5.1	Aim and structure	53
5.2	What is a performance contract?	53
5.3	Conditions for effective performance contracts	54
5.4	Empirical evidence on the potential of performance contracts for policing	57
5.5	Conclusions	58
6	The proposed performance contracts	63
6.1	Aim and structure	63
6.2	Brief history and plans for the future	63
6.3	The design of the performance contracts	64
6.4	Conclusions	65
7	Anticipating gaming of the performance evaluation	67
7.1	Aim and structure	67
7.2	Objective performance measures and the tradeoff between incentives and gaming	67
7.3	Shifting from objective to subjective performance evaluation	73
7.4	Conclusions	75
8	Effectiveness of the proposed performance contracts	77
8.1	Aim and structure	77
8.2	Assessing the proposed method of performance evaluation	77
8.3	Assessing collection of data	82
8.4	Assessing the proposed design of the performance rewards	83
8.5	Conclusions	87
9	Brief overview of policies in the United Kingdom and Australia	91
9.1	Aim and structure	91
9.2	Policies in England and Wales	91
9.3	Policies in Australia	94
9.4	What can we learn from these countries?	95
10	Conclusions	99
10.1	Aim and structure	99

10.2	What goals might performance contracts serve?	99
10.3	Are the proposed performance contracts likely to be effective?	100
10.4	What are options for policy?	102
	Abstract	107
	References	109
	Annex 1. The measures for the system of performance rewards	117
	Annex 2. People consulted	119

Preface

The regional police forces are seeing their budgets grow in a time of cutbacks in public expenditures. Regional police forces have a great degree of autonomy in choosing how to allocate these funds. The great degree of discretion makes regional police forces responsible for thinking about what they want to achieve and the manner in which they will attain their goals. Performance contracts could be instrumental in making these choices explicit. As such, they could contribute to improved police accountability and incentives.

The nature of police work makes the application of performance contracts highly challenging. Many questions still need to be answered. Can performance measures be selected that reflect the quality and variety of police work? Can targets be set at the right level? Are financial incentives a good way of rewarding efforts? These issues have been discussed in a range of research studies and policy documents - particularly by policy makers and analysts in Anglo-Saxon countries. In this study, Ben Vollaard reviews some of the existing literature in this area and applies it to the current debate on performance contracts for the police forces.

We are indebted to the people listed in the annex, who were interviewed or consulted in conducting the analysis. Several people kindly took the time to review a draft of the report, and provided many useful suggestions on how to structure the analysis and how to present the results. Special thanks go to Frits Vlek from the 'Commissie Politie en Wetenschap', Caspar Wiebrens from the 'Parket Generaal' of the Prosecution Council, Oskar Huurdeman and Esther de Kleuver from the Interior Department, and CPB colleagues Pierre Koning, Taco van Hoek, Marcel Canoy, Richard Venniker and Marc Pomp. We owe thanks to Sam Loeb for linguistic advice and to Simone Pailer for the lay out of the report. We would like to thank the Department of Economic Affairs and the Institute for Humane Studies for their financial support.

Henk Don, director

Executive summary

In the beginning of 2003, the Secretary of the Interior, the Minister of Justice and police administrators signed the 'National Police Covenant'. The covenant states the way by which the police forces will contribute to the national goal of reducing crime and disorderly behaviour by 20 to 25 percent over the next three years. As part of the covenant, performance contracts will be signed with each individual regional police force by no later than July 2003. The contracts establish financial rewards for meeting a number of quantitative performance targets. The targets are related to output (fines, charges), citizen satisfaction, and internal performance (including sickness absence and telephone response times). The expectation is that the performance contracts will improve police accountability and incentives.

Concerns about police performance

In accordance with the government's ambitions, the scarce empirical evidence indicates that there is scope for improvement in police accountability and incentives. Data from the prosecution council show great differences in performance between the forces. Moreover, some poorly performing forces are not catching up. These findings indicate that there is room for a more lively exchange of best practices between the forces.

Recent empirical studies provide indications that the overall performance of the police has been declining. Crime clear-up rates have been on the decrease over the last 20 years, whereas they have been stable or improving in other countries. The number of solved crimes per officer (weighted with the respective penalty) has decreased over the last five years. When cleaning police data from red light and speed camera tickets, the overall productivity of the police appears to be on the decrease since 1990. These findings indicate that there is also room for a more systematic improvement of policing strategies through carefully evaluated experiments.

However, based on the available data, we are not able to draw hard conclusions on the background of the decline in measured police performance. Unfortunately, many of the rich sources of administrative data of the forces have not been opened up yet. Consequently, little is publicly known about the way police budgets have been spent or the results of police work. Is the police unable to cope with changes in crime patterns and disorderly behaviour? Or are police officers choked by an ever-increasing administrative burden and a lack of upstream capacity in the criminal justice system? The lack of information on police efficiency and effectiveness alone is an indication of the low importance attached to understanding and improving police performance.

Performance contracts as seemingly attractive policy instrument

At first sight, entering into performance contracts with the police forces seems to be an attractive way of improving their accountability and incentives. The rewards for meeting the performance targets support the motivation to improve. The rewards compensate the forces for the time and effort spent on evaluating operations and motivating employees to share and implement best practices. Moreover, the rewards acknowledge the efforts of police officers and managers who have to overcome (bureaucratic) barriers to change and the continuous, strong pressure to respond to the issues of the day.

Proposed performance contracts not likely to be very effective

A closer look at the proposed performance contracts learns that they are not likely to be very effective in achieving verifiable improvements in police performance. We base this conclusion on the following findings:

- *Fixing desirable improvements into a contract invites the forces to ‘meet the numbers’ at the cost of quality of police work.* The performance contracts allow for several ways of achieving the targets, many of which are not in line with the government’s objectives. If the forces are expected to react to the (financial) incentives, then paying off for behaviour other than what the government is seeking undermines the effectiveness of performance contracts. The contracts reward the forces for substituting away from tasks that do not affect the performance measures to tasks that do (at the possible cost of pro-active crime prevention). Within activities that affect the measures, the contracts also pay off for lowering the quality of output (frivolous fining, picking the easiest criminal charges, producing charges with questionable evidence). These are all easy and therefore tempting ways of meeting the targets.
- *Fixing levels of output conflicts with the need to respond to changing regional circumstances.* By putting the desired number of fines and charges in a contract, they are made independent of changes in the nature and frequency of criminal and disorderly behaviour. The targets limit the responsiveness of regional police forces to changes in their environment. Therefore, the demands of the performance contracts and reality may well start to deviate. Clearly, this is not in the interest of effective policing.
- *The government does not have the data to set the performance targets at the right level.* In the absence of any firm knowledge about the means by which the targets can be achieved, each target included in the performance contract is likely to be either out of reach or too little ambitious. Detailed information on business processes and regional circumstances is lacking. Back-of-the-envelope calculations indicate that the targets for the number fines and charges are little ambitious - certainly given the future increase in the number of police officers. The goals for customer satisfaction may also be little ambitious since they are based on the own historical

performance ('perform at least as good as you did in the past'). If the targets are set too low, the force may stop their efforts as soon as the target is reached ('satisficing behaviour').

- *The performance targets conflict with the wish for flexibility at the national level.* Frequently, the government places new demands on the police forces. These changing policy priorities may well make the performance targets obsolete. When anticipating on future changes, the forces would have little incentive to meet the current targets. Furthermore, negative experiences with some of the performance measures may prompt the government to change the terms of contract. Such changes are not unthinkable given the inexperience with this policy instrument. Altering the method of measuring and rewarding performance along the way does not support the forces' incentives either.
- *The financial rewards could provoke undesirable biases in police work.* Awarding the financial reward requires a yes/no decision based on weighing multiple performance measures. It is hard to replace subjective judgement with mechanical rules, however. The government does not have the data to choose the appropriate weights. When the weights included in the decision rules are picked wrong, undesirable biases in police work may result. The forces can be expected to look for the target that can be achieved with the least effort and the greatest reward.
- *The financial rewards conflict with the lack of consumer choice.* By not increasing the budgets of forces that did not sufficiently improve their performance, citizens pay twice for poor police management within their region. In contrast with schools and hospitals, citizens cannot switch to a better performing force other than by moving somewhere else.

Some of the shortcomings can be alleviated...

The performance contracts could be made more effective by making changes. Some of the options are:

- *Search for better performance indicators.* Better measures could limit the side effects and lower the conflict with flexibility at the regional level. Of all measures included in the contract, the output measures number of fines and number of charges most strongly invite the forces to just 'meet the numbers'. They also limit the discretion of the forces most strongly. An option is not to include the number of fines in the contract. The charges can be weighted with their respective penalty. Then the forces no longer have the incentive to focus on the easiest cases.
- *Tie financial incentives to police managers' remuneration rather than the forces' budgets.* That way, customers of police services will no longer pay twice for poor police performance.
- *Gather better data to improve target setting.* Better data on how the forces spend their resources and what outputs are produced are necessary to make a better informed guess on what the targets should be. To that end, the rich sources of administrative data of the forces would need to be opened up.

... but there remain serious drawbacks to the use of performance contracts

Also with these changes, there remain serious drawbacks to the use of performance contracts. The need to react to changing regional and national circumstances will always make it hard to contract police output in advance. Setting targets at the right level requires a level of knowledge about the specifics of police work and regional circumstances that is unlikely to be present at the national level. Furthermore, the quality and variety of police work is just hard to capture in quantitative performance measures. Therefore, the performance measures will always allow for ways of meeting the targets that are not in line with the government's objectives.

Towards an alternative approach: organizing accountability to customers and peers

Considering the serious drawbacks related to the contractual approach, it is worthwhile to look across the borders for alternative ways of achieving the government's objectives. Australia and the United Kingdom are two countries that have been working on greater police accountability and incentives for a long time. Interestingly, in these countries experts have been closely involved in the design of policies towards the police. Neither Australia nor the United Kingdom have opted for performance contracts with financial incentives for the forces. Instead their prime focus is on benchmarking the police forces. Many years have been invested in improving the reliability of police data and methods of benchmarking forces in a fair and meaningful way.

A similar approach could be worked out for the Dutch police. In that case, the role of the national government is not to put down desired changes in a contract but to organize accountability to customers and peers. The outcomes of evaluations by customers and peers from other forces can be an input into decisions about promotion and remuneration of a force's management. Customers and peers are in the best position to assess performance - and to provide ideas for improvement. Customers not only include citizens currently surveyed through the Citizen Police Monitor (PMB), but also businesses and public prosecutors. Managed peer reviews are a necessary mean to conduct a thorough and comprehensive assessment of policing strategies. Through peer reviews, the management of each police force is forced to develop and present a coherent picture of the problems in their region, their strategy to tackle these problems, and the results relative to other forces. This approach of 'organized accountability' could contribute to a culture within the police of systematically following, evaluating and improving performance.

Samenvatting (summary in Dutch)

Begin 2003 tekenden de Minister van Binnenlandse Zaken en Koninkrijksrelaties, de Minister van Justitie en de korpsbeheerders het 'Nationaal Politie Convenant'. Het convenant geeft aan welke bijdrage de politie zal leveren aan de nationale doelstelling een vermindering te bereiken in crimineel en orde verstorend gedrag van 20 to 25 procent over de komende drie jaar. Als onderdeel van het convenant sluit de overheid vóór juli 2003 prestatiecontracten af met elk regionaal politiekorps. In de contracten staat dat de korpsen een financiële beloning zullen ontvangen als zij een aantal kwantitatieve doelstellingen behalen. De afspraken betreffen productie (bekeuringen, zaken naar Openbaar Ministerie), tevredenheid van burgers en interne organisatie (onder andere ziekteverzuim en snelheid van telefonische beantwoording). De verwachting is dat de prestatiecontracten zullen bijdragen aan de prestatieprikkels voor en verantwoording door de politiekorpsen.

Zorgen over politiestatistiek

De resultaten van het weinige beschikbare empirische onderzoek geven aan dat er ruimte is voor verbetering in prestatieprikkels en verantwoording. Gegevens van het Openbaar Ministerie laten grote verschillen zien in de prestaties van de korpsen. Bovendien halen sommige minder presterende korpsen hun achterstand niet in. Deze bevindingen vormen een aanwijzing dat er ruimte bestaat voor een meer levendige uitwisseling van *best practices* tussen de korpsen.

Ook laten recente empirische studies een daling zien in de prestatie van de politie. Over de afgelopen 20 jaar zijn de ophelderingspercentages voortdurend gedaald, terwijl ze stabiel waren of zelfs verbeterden in andere landen. Het aantal opgehelderde misdrijven per agent (gewogen met de bijbehorende straf) is gedaald over de laatste vijf jaar. Wanneer politiegegevens geschoond worden voor flietspaal-boetes blijkt de productiviteit van de politie sinds 1990 te dalen. Deze resultaten geven een indicatie dat er ruimte is voor een meer systematische aanpak van het verbeteren van de prestatie met behulp van zorgvuldig geëvalueerde experimenten.

Op basis van de beschikbare gegevens kunnen we echter geen harde conclusies trekken over de precieze achtergrond van de gemeten daling in politiestatistiek. Helaas zijn de rijke bronnen van administratieve gegevens van de korpsen nog niet goed toegankelijk. Hierdoor is publiekelijk weinig bekend over de manier waarop de budgetten worden besteed of over de resultaten van politiewerk. Is de politie niet in staat veranderingen in criminaliteit en orde verstorend gedrag bij te houden? Of zijn politieagenten verstikt in een voortdurend groeiende administratieve rompslomp en een gebrek aan capaciteit hoger in de strafrechtketen? Het gebrek aan informatie over efficiëntie en effectiviteit van de politie is zelf al een indicatie van het lage belang dat gehecht wordt aan het begrijpen en verbeteren van de prestatie van de politie.

Prestatiecontracten als een schijnbaar aantrekkelijk beleidsinstrument

Op het eerste gezicht lijkt het afsluiten van prestatiecontracten met de regiokorpsen een aantrekkelijke manier om de verantwoording en prikkels te versterken. De prestatiebeloning ondersteunt de motivatie om te verbeteren. De beloning compenseert de korpsen voor de tijd en inspanningen die gaan zitten in het evalueren van hun optreden en het motiveren van agenten om *best practices* te delen en te implementeren. Bovendien vormt de beloning een erkenning van de inspanningen van agenten en managers die nodig zijn om de (bureaucratische) barrières tegen verandering en de waan van de dag te doorbreken.

Voorgestelde prestatiecontracten zijn waarschijnlijk weinig effectief

Een nadere analyse van de voorgestelde prestatiecontracten geeft aan dat zij waarschijnlijk weinig effectief zijn in het bereiken van aantoonbare verbeteringen van de veiligheid. Wij baseren deze conclusie op de volgende bevindingen:

- *Het vastleggen van gewenste verbeteringen in een contract nodigt uit tot een gerichtheid op cijfertjes die ten koste gaat van de kwaliteit van politiewerk.* De contracten laten ruimte voor vele manieren om de prestatieafspraken te behalen. Niet alle manieren zijn in lijn met de doelstellingen van de overheid. Als van de korpsen verwacht wordt dat zij reageren op (financiële) prikkels, dan ondermijnt het belonen van zulk gedrag de effectiviteit van prestatiecontracten. De contracten belonen de korpsen voor het substitueren van taken die niet tot uitdrukking komen in de prestatieafspraken naar taken die dat wel doen (mogelijkerwijs ten koste van actieve misdaadpreventie). Binnen activiteiten die de prestatieafspraken beïnvloeden, belonen de contracten ook het verlagen van de kwaliteit van de productie (pietluttig bekeuren, de eenvoudigste zaken oppakken en het doorsturen van zaken met twijfelachtig bewijs). Dit zijn allemaal eenvoudige en daarom verleidelijke manieren om de prestatieafspraken te behalen.
- *Het vastleggen van te leveren productie rijmt niet met de noodzaak om te reageren op veranderende regionale omstandigheden.* Door het gewenste aantal bekeuringen en zaken in een contract vast te leggen worden deze grootheden onafhankelijk gemaakt van de aard en het niveau van crimineel en orde verstoring gedrag. De prestatieafspraken beperken dus de mogelijkheid van de korpsen om te reageren op veranderingen in hun omgeving. Hierdoor kunnen de eisen van de prestatiecontracten en de werkelijkheid uit elkaar gaan lopen. Dit is natuurlijk niet in het belang van effectief politiewerk.
- *De overheid kan vanwege een gebrek aan informatie het ambitieniveau van de afspraken niet goed inschatten.* Zonder goede informatie over de middelen waarmee de prestatieafspraken zijn te bereiken, is iedere doelstelling in het prestatiecontract waarschijnlijk óf buiten bereik óf te weinig ambitieus. Er is een gebrek aan gedetailleerde informatie over bedrijfsprocessen en regionale omstandigheden. Berekeningen op de achterkant van een sigarendoos laten zien dat de doelstellingen voor het aantal bekeuringen en zaken waarschijnlijk weinig ambitieus zijn -

zeker gezien de toekomstige groei in het aantal politieagenten. De doelstellingen voor klanttevredenheid lijken ook weinig ambitieus, omdat ze gebaseerd zijn op de eigen prestatie in het verleden ('presteer tenminste zo goed als in het verleden'). Als de afspraken weinig ambitieus zijn, kunnen de korpsen hun inspanningen verlagen zodra de doelstelling is behaald.

- *De prestatieafspraken conflicteren met de behoefte aan flexibiliteit op nationaal niveau.* De overheid stelt regelmatig nieuwe eisen aan de politiekorpsen. Deze veranderende beleidsprioriteiten kunnen de gemaakte afspraken betekenisloos maken. Als de korpsen anticiperen op veranderingen in de toekomst hebben ze weinig prikkel om nu de afspraken te behalen. Bovendien kunnen negatieve ervaringen met prestatie-indicatoren aanleiding geven tot een wijziging in het contract. Veranderingen in het meten en belonen van de prestatie zijn niet ondenkbaar gegeven het gebrek aan ervaring met dit beleidsinstrument. Het tussentijds veranderen van het contract verzwakt ook de prestatieprikkels.
- *De financiële prestatiebeloning nodigt uit tot ongewenste verschuivingen in politiewerk.* Het toekennen van de prestatiebeloning vereist een ja/nee beslissing gebaseerd op het wegen van verschillende prestatie maatstaven. Het is echter moeilijk om een subjectieve beoordeling te vervangen door simpele beslisregels. De overheid heeft niet de gegevens om de juiste gewichten te kiezen. Verkeerd gekozen gewichten kunnen leiden tot ongewenste verschuivingen in politiewerk. Van de korpsen kan immers worden verwacht dat zij zich richten op doelstellingen die met de minste inspanning en tegen de hoogste beloning zijn te realiseren.
- *De financiële prestatiebeloning conflicteert met het gebrek aan consumentenkeuze.* Door het budget van minder goed presterende korpsen niet te verhogen betalen burgers een dubbele rekening voor slecht politiemangement in hun regio. In tegenstelling tot scholen en ziekenhuizen kunnen burgers niet kiezen voor een beter politiekorps anders dan door te verhuizen.

Sommige tekortkomingen zijn te verlichten...

De effectiviteit van de prestatiecontracten is te verhogen door het maken van een aantal aanpassingen. Hiervoor zijn verschillende opties:

- *Zoek naar betere prestatie-indicatoren.* Betere prestatie-indicatoren kunnen de negatieve bijwerkingen verkleinen en tegemoet komen aan de behoefte aan flexibiliteit. Van alle maatstaven in het contract nodigen het aantal boetes en het aantal aan het Openbaar Ministerie te leveren zaken het sterkst uit tot ongewenste gedragseffecten. Ook beperken zij de beleidsvrijheid van de korpsen het meest. Een optie is het aantal boetes te schrappen als prestatie maatstaf. Ook is het mogelijk de opgehelderde zaken te wegen met de zwaarte van het delict. Dan heeft de politie niet langer een prikkel zich op de makkelijkste zaken te richten.
- *Verleg de financiële prikkels van het budget naar de beloning van managers.* Op deze wijze betalen de klanten van de politie niet langer een dubbele rekening voor matige politieprestaties in hun regio.

- *Verzamel betere gegevens om het ambitieniveau van de prestatieafspraken beter in te schatten.*
Hiervoor zijn betere gegevens over bedrijfsprocessen nodig. Het toegankelijk maken van de rijke bronnen van administratieve gegevens van de korpsen is hiervoor noodzakelijk.

... maar de prestatiecontracten blijven belangrijke negatieve bijwerkingen houden

Ook met deze aanpassingen blijft het gebruik van prestatiecontracten belangrijke negatieve bijwerkingen houden. De noodzaak om te reageren op veranderende regionale en nationale omstandigheden is moeilijk te verenigen met de wens prestaties vooraf vast te leggen. Het inschatten van het ambitieniveau van de doelstellingen vereist een niveau van kennis over bedrijfsprocessen en regionale omstandigheden dat niet goed is te realiseren op nationaal niveau. Bovendien is de kwaliteit en variëteit van politiewerk gewoon niet goed te vangen in een aantal prestatie-indicatoren. Hierdoor zal er altijd ruimte blijven bestaan voor gedragseffecten die niet in lijn zijn met de doelstellingen van de overheid.

Naar een alternatieve benadering: het organiseren van verantwoording aan klanten en collega's

Gezien de tekortkomingen van de contractuele benadering is het waardevol om over de grenzen te kijken naar alternatieve wijzen om de doelstellingen van de overheid te realiseren. Australië en Engeland zijn twee landen die al lange tijd werken aan betere verantwoording en prikkels voor de politie. Het is interessant te zien dat in deze landen experts nauw betrokken zijn bij de vormgeving van het politiebeleid. Australië noch Engeland hebben gekozen voor prestatiecontracten met financiële prikkels voor de korpsen. Zij focussen op het benchmarken van korpsen. Vele jaren zijn geïnvesteerd in het verbeteren van de betrouwbaarheid van gegevens en van methoden om de korpsen op een eerlijke en betekenisvolle manier te vergelijken.

Een soortgelijke benadering is uit te werken voor de Nederlandse politie. Volgens deze aanpak is de rol van de nationale overheid niet om gewenste verbeteringen in een contract vast te leggen, maar om verantwoording aan klanten en collega-korpsen te organiseren. Het management van de korpsen kan worden afgerekend op basis van de uitkomsten van de evaluatie door klanten en collega-korpsen (*peer review*). Klanten en collega's zijn in de beste positie om politieprestaties te evalueren - en om ideeën voor verbetering aan te dragen. Klanten zijn niet alleen burgers die nu worden geënquêteerd door de Politie Monitor Bevolking, maar ook bedrijven en officieren van justitie. Intervisie tussen korpsen is noodzakelijk voor een diepgravende en brede evaluatie van de prestaties van een korps. De intervisie tussen korpsen dwingt het management van ieder korps om een coherent beeld te ontwikkelen van de veiligheidssituatie in de regio, de strategie om de problemen aan te pakken en de behaalde resultaten in vergelijking met andere korpsen. Deze aanpak van 'georganiseerde verantwoording' kan bijdragen aan de ontwikkeling van een cultuur binnen de politie die zich kenmerkt door het systematisch volgen, evalueren en verbeteren van de prestatie.

1 Introduction

1.1 Rationale

Of all public tasks, Dutch citizens give fighting crime and maintaining public order highest priority (SCP, 2001). However, the performance of police services falls well short of expectations. In no other industrialized country do citizens rate their satisfaction with the work of the police lower than in the Netherlands (SCP, 2002, p. 666). The typical response to calls for better policing has been an increase in police funding. The number of citizens served by each police officer has been brought down from 390 in 1980 to 320 in 2002, a decrease of almost 20 percent. By 2010, this figure will be as low as 306, a further decrease of 4 percent (BZK and Justitie, 2003, p. 2).

However, simply boosting general resource levels to improve safety is no longer seen as sufficient. Public attention is slowly being refocused from the number of officers to what they actually do and the effects of their daily work. So far little use has been made of the rich data sources owned by the police forces (one of the exceptions is PKP, 1988). Consequently, little is publicly known about the way police budgets have been spent or the efficiency and effectiveness of the police. Are the regional police forces getting the most of their yearly budget of EUR 2.7 billion? What can the public expect from a further increase in the number of police officers? Are the streets going to be safer? And if so, how do we know whether the police made the difference? Recent claims that the efficiency of the police is declining dramatically present more questions than answers.¹ To what extent do capacity restraints in the criminal justice system put a check on police work? Are administrative burdens keeping the officers in the police station?

To address concerns about accountability and police performance, the Secretary of the Interior is about to introduce performance contracts for the 25 regional police forces. Measures that indicate the performance of the police have been identified, targets and financial incentives for meeting those targets have been set. The regional performance contracts will be based on the national police covenant that was signed on February 15, 2003. The performance contracts are thought to contribute to greater accountability and a shift towards a results-oriented police organization (BZK, 2002c, p. 72).

¹ See Smit (2000) and Tak and Fiselier (2002) for international comparisons and Wiebrens (2002b) for trends within the Netherlands.

Critical observers claim that these contracts are like a Trojan horse.² The publicly announced targets and the accompanying flow of performance indicators are appealing in the sense that they create the appearance of control over the police. However, the actual use of such contracts could lead to various undesirable side effects. For example, a strong focus on performance indicators could result in a bias towards objectives that can be quantified at the expense of objectives that cannot so easily be quantified. On the basis of such arguments, sceptics claim that the remedy is worse than the disease.

Uncertainty about the behavioural effects of performance contracts is not unique to the case of Dutch police forces. Other sectors and countries are facing similar problems. For instance, policy makers and analysts in the United Kingdom have spent a considerable amount of time on improving the accountability of their police (see Drake and Simper, 2001 for an overview). Similarly, Australia has built up a body of highly relevant literature on this topic (see Carrington et al., 1997 for instance). Especially given the experimental stage of the Dutch system of performance contracting, it is important that we learn from the insights and experiences that have been built up elsewhere.

1.2 Research question

The following research question is central to this study:

Is it likely that performance contracts for the regional police forces will contribute to realization of the Interior Department's objectives for regional police work?

It is important to note that we focus on regional police work, not on supra-regional issues such as fighting international crime syndicates. The latter activities are often organized at the national or international level. We will further discuss the scope of this study in Section 1.8.

1.3 Contribution to policy making

We aim to support decision-making at the Department of the Interior by critically reviewing the potential contribution of performance contracts. At present, the performance contracts are in the making. They will be introduced by no later than June 30, 2003. As argued above, an application of the insights and experiences documented in the literature to this practical problem could help the Department appropriately implement this new policy instrument.

² See Schoof (2002) for an interview with Frits Vlek on the Dutch initiatives. For a critical review of British and Australian experiences see FitzGerald et al. (2002) and Dupont (2001)

Clearly, there is a particular need for input in the months to come. This study is meant to be a timely contribution to policy making. True to the mission of the CPB Netherlands Bureau for Economic Policy Analysis, we will clarify the choices and tradeoffs that are to be made rather than present politically coloured answers.

1.4 The Interior Department as principal

We take the perspective of the Department of the Interior. This Department can be seen as principal of the 25 regional police forces because of its responsibility for the efficiency and effectiveness of the police (BZK, 2002c, p. 66 and Police Act, article 53a). Over the last years, the role of the Department as principal at the national level has been strengthened.³ However, the position of the Department is different from other principals managing independent bodies.⁴ Each force is embedded in a regional administrative structure that includes mayors and public prosecutors, the respective authorities for maintaining public order and emergency assistance and law enforcement (see the box below). Actual goals and priorities therefore tend to set at the regional and the local level.⁵ For this reason, we use a broad definition of a regional police force that includes the force but also the regional authorities (mayors and public prosecutors). They together are responsible for realizing the objectives of the Department within their region. When discussing the design of performance contracts, we explicitly account for the Department's unique relation to the forces.

³ TK 26345, Nr. 79 (2002): "During this administration the responsibility of the Secretary of the Interior for central management of the police, as well as actual policing has been supported with wider powers. (...) The powers to manage the regional police forces at the national level are concentrated within the Department of the Interior." Translated from: "In deze kabinetsperiode is de verantwoordelijkheid van de minister van BZK voor het centraal beheer van de Nederlandse politie, alsmede voor de politiezorg nader ondersteund met bevoegdheden. Zo (...) zijn de bevoegdheden op rijksniveau met betrekking tot het beheer van de regionale politiekorpsen bij BZK geconcentreerd."

⁴ Each force has the legal status of independent body ('Zelfstandig Bestuursorgaan', ZBO). A ZBO has a high degree of autonomy in performing its public task. Other examples of ZBOs are the Social Benefit Administration (UWV), the Central Bank (DNB) and the Department of Motor Vehicles (CBR).

⁵ BZK and Justitie (1998): "The differences in safety between the 25 regions (but also within these regions) make it self evident and desirable that the forces set their own priorities (...) to achieve the objectives set at the regional level." Translated from: "De onderlinge verschillen in veiligheid tussen de 25 regio's (maar ook daarbinnen) maken het vanzelfsprekend en wenselijk dat de korpsen (...) zelf de accenten kunnen leggen voor het bereiken van regionaal bepaalde (veiligheids-) doelstellingen."

Management at the regional level

At the regional level, each force is managed by the mayor of the main municipality in the region ('korsbeheerder'), here forth 'mayor in command'. He is accountable to the regional council ('regionaal college'), which he also presides. The council consists of all mayors in the region and the chief public prosecutor ('hoofdofficier van justitie'). The regional council makes all final decisions on the budget, annual account and the regional policy plan ('regionaal beleidsplan'). The regional policy plan includes the objectives and priorities of the force. The mayor in command drafts the budget, annual account and the policy plan together with the chief public prosecutor. The (elected) local councils can advise their mayor on issues that are to be decided by the regional council. The chief of police ('korschef') assists the mayor in command with daily management of the force. He has no formal powers. He meets the mayor in command and the chief public prosecutor on a regular basis to discuss management issues ('regionaal driehoeksoverleg'). There are similar regular meetings at the local level between the mayor, the public prosecutor and local chief of police ('locaal driehoeksoverleg'). The most striking feature of the regional institutional structure is that all players have multiple roles. Clearly, one of the key players at the regional level is the mayor in command. He combines many roles in one person. He is not only involved in drafting the regional police plan, but also in accepting it as chairman of the regional council. Moreover, he supervises proper execution of the plan.

1.5 Approach

In this study, we apply insights from the literature on the theory and practice of performance contracts to the problem of assessing the value of such contracts for management of the regional police forces. The analysis is qualitative of nature because of its focus on the optimal design of an institution that is still in the making.

Economic theory provides a useful framework for studying tradeoffs in the design of performance contracts for police forces. The principal-agent model - developed in the economics literature regarding incentives in organizations - provides a meaningful way of analysing the interaction between the Department of the Interior and the police forces.⁶ A substantial part of current research on the design of performance contracts is conducted by economists or based on economic theory.

Most of the foreign experiences documented in the literature that we review are based on policies in the United Kingdom and Australia. Several years ago, both countries started a comprehensive review of their police management systems. Since that time, both the United Kingdom and Australia have worked hard to improve accountability of the police. Consequently, they are ahead of the Netherlands when it comes to the evaluation of police performance and

⁶ So far, the principal-agent theory has not been used extensively in the area of police studies. Exception known to the author is Prendergast (2001). In their review of the literature on the economics of law enforcement, Polinsky and Shavell (2000, section 19) put it as line of inquiry that merits further development.

design of incentives for the police. There could be much to learn from experiences in these countries.

In order to maintain a practical understanding of the issues, we conducted open interviews with people who are involved in policy making and research in this area (see Annex 2 for a list of consulted people).

1.6 Structure

To assess whether performance contracts for the police forces are likely to have a positive effect, we structure the paper along the following three steps:

Step I. What goals might performance contracts serve?

As a first step, we discuss why the Department of Interior would need performance contracts. To that end, we analyse whether the Department's current policies - without performance contracts - stimulate the forces to work towards their objectives. The objectives for regional police work are defined in Chapter 2. We continue with an analysis of the forces' incentives to realize these objectives in Chapter 3 (before the introduction of performance contracts). In Chapter 4, we assess whether the current empirical evidence is in line with our findings in Chapter 3. Finally, we discuss whether performance contracts are a logical policy instrument to address the need for change in Chapter 5.

Step II. Are the proposed performance contracts likely to be effective?

After defining the goals of performance contracts, step II assesses whether the proposed performance contracts are likely to be effective. In Chapter 6 we provide a description of what is known at this time about the design of the proposed performance contracts. In Chapter 7 we discuss the possible behavioural effects of these contracts. This chapter is the core of the analysis. The conclusions of Chapter 8 provide input for the third step in the analysis: how could the design be improved?

Step III. What is the value of performance contracts for the police?

In Chapter 9, we will provide a brief overview of the policies that are being pursued in Australia and the United Kingdom. This look across the borders is meant to provide ideas for alternative ways of achieving the Department's objectives. Based on the analysis in the foregoing chapters and the lessons from Chapter 9, we draw conclusions about the value of performance contracts for the police forces in Chapter 10.

1.7 Defining the object of analysis

Here we use a descriptive definition of police services. The table below provides an overview of the activities in which the police forces engage themselves. Later on in the report, we will discuss the objectives that connect these activities to a common purpose. We do not include privately provided police services in the analysis (such as private investigators, private security surveillance, etc.). In the analysis, we take the public nature of the police as given: we do not go into the discussion of public versus private provision (see Van der Vijver et al., 2001).

Table 1.1 Police tasks according to the Activities Code Table (ANP)

Task	Definition
Community policing	Provision of basic police services by a police unit that works exclusively in a certain geographic working area.
Emergency assistance	Providing services to people in acute emergencies on a 24/7 basis.
Criminal investigations	Investigation of crimes directed at finding the offender and/or goods. This includes all activities related to support to the public prosecutor.
Problem-oriented policing	Activities that are directed at a specific item, including large events, traffic, and moral law.
Customer-oriented policing	Activities at the police station related to in-person or telephone contact with the public. This includes reporting a crime.
Special police tasks	These activities are exclusively related to protection of persons/ goods, execution of penalties and border patrol.

Note: The Activiteitencodetabel Nederlandse Politie (ANP) includes an exhaustive list of police tasks to provide a uniform framework for activity measurement.

1.8 Scope

To keep the analysis tractable, we had to decide which issues are truly essential to the research question and which issues could be left outside the scope of this study. There are four major areas that would justify separate studies.

Interaction within the forces

We focus on the interaction between the Department of the Interior and the regional police forces. As discussed above, there are major policy initiatives at this level - and many unanswered questions. An analysis of incentives within the forces is beyond the scope of this study (the interested reader is referred to Terpstra, 2002).

Supra-regional police work

We focus on police work that has an exclusive regional dimension. We do not go into issues that are beyond the control of a single police force, such as fighting organized crime (the so-called ‘bovenregionale of landelijke politiediensten’) or infrastructure issues such as shared information and communication technology (ICT). Indeed there are major concerns in this area, particularly about coordination between the forces (Algemene Rekenkamer, 2003b concludes that exchange of information for criminal investigations is in the same sorry state as in 1985). Some of these concerns, such as the exchange of information between forces, directly affect issues discussed in this report. After all, without a good facilitating role of the Department, exchange of information within and between the forces will be hindered.⁷ However, there are different causes underlying a role for the Department in supra-regional police work (such as externalities and coordination failure) than in the case of exclusively regional police work. Because of the different rationale for intervention by the Department of the Interior, these issues require a different institutional framework (nationally organized units, a centralized ICT organization, etc.). Therefore, an analysis of the role of the Department in supra-regional police work would require a separate study (see for example Kleemans, Brienens and Van de Bunt, 2002).

Changing the organizational framework

Some observers claim that the current organizational framework of the police does not foster a culture of accountability and incentives to improve performance. There are many people involved in making policy for the police, but not many of them bear responsibility for the outcomes. Vague definitions of responsibilities may hinder accountability (Algemene Rekenkamer, 2003a, p. 18, 19). There is little democratic control of the police forces (Muller, 2002). There is no good distinction between policy making and evaluation (ibid.). The mayors in command have the role of policy maker and auditor (see Section 1.3). The Police Inspection does not have a strong role (Algemene Rekenkamer, 2001). The public prosecutor has only one voice in the regional management council against many mayors. Consequently, there may be too little attention to performance in the area of criminal investigations. Clearly, the organizational framework of the police has effects on performance. We see the organizational framework as a separate issue however. In this study, we take the organizational framework as given and focus on the value of performance contracts.

⁷ BZK and Justitie (1998 p. 33): “... often there has been little attention to cooperation between the forces. That has resulted in a patch work of information and management infrastructures.” Translated from: ‘... is er vaak minder aandacht geweest voor onder meer samenwerking tussen de korpsen onderling. Dat heeft ertoe geleid dat bijvoorbeeld de informatiehuishouding en de bijbehorende bedrijfsvoering op verschillende manieren zijn ingericht.’

Behavioural responses outside the economic model

The analysis is based on the economic model of behaviour assuming rational actors that maximize their utility subject to a number of constraints. Behavioural responses that cannot be reduced to this model are outside the scope of this study. We are convinced that the economics literature provides a meaningful framework for studying the behavioural effects of performance contracts. Our comparative advantage is in economics. We leave it to people specialized in other disciplines to complement our analysis. For example, see Eisenberger and Cameron (1996) for a highly instructive review of the effects of incentives on motivation from a psychological perspective.

Step 1. What goals might performance contracts serve?

Is there a need for change in the Department's policies towards the forces? And if so, are performance contracts likely to foresee in the need? These are the questions that we address in the first step of the analysis. We define the objectives for regional police work in Chapter 2. Then, we go on to analyse the forces' incentives to realize these objectives in Chapter 3 (before the introduction of performance contracts). In Chapter 4, we assess whether the current empirical evidence is in line with our findings in Chapter 3. Finally, we discuss whether performance contracts are a logical policy instrument to address the need for change in Chapter 5.

2 The government's objectives for regional police work

2.1 Aim and structure

What are the objectives for regional police work? By first defining a clear set of objectives, we can analyse the Department's policies aimed at realizing these objectives in later chapters. We will argue that the desired discretion of the police forces implies a concern with the overall performance of the forces rather than a detailed prioritization of goals for regional police work. The forces should provide efficient, effective, fair and decent policing within the framework of broadly defined outcomes of regional police work.

We take a number of steps before arriving at this conclusion. First, we discuss the implications of translating the tasks of the police laid down by law into objectives in Section 2.2. Then, we discuss the Department's interest in providing great discretion to the regional level in Section 2.3. Clearly, the unique relation between the Department and the regional level has implications for the way the objectives are to be defined. In Section 2.4, we discuss these implications and provide a possible definition of the objectives. Section 2.6 concludes.

2.2 Translating tasks into objectives

The objectives for the regional police forces are based on the task of the police described in the Police Act of 1993, Article 2:

*It is the task of the police to maintain the legal order and to provide help to those who need it - in subordination to the proper authorities and in accordance with the rule of law.*⁸

More specifically, the task areas of the police are (1) maintaining public order (article 12 Police Act 1993), (2) enforcing the law (article 12 Police Act 1993) - this task area includes enforcement of a number of special laws including the Environmental Management Act, Immigration Law, Licensing Act and Traffic Act - and (3) providing emergency assistance (article 2 Police Act 1993).

The general description of tasks needs to be translated into objectives. The actual wording of the translation is essential for the range of activities the police will get involved in. By translating

⁸ Translation from Politiewet 1993: "De politie heeft tot taak in ondergeschiktheid aan het bevoegd gezag en in overeenstemming met de geldende rechtsregels te zorgen voor de daadwerkelijke handhaving van de rechtsorde en het verlenen van hulp aan hen die deze behoeven."

tasks into objectives, police work is extended beyond the tasks described in the law. For example, a burglary prevention campaign is not written into the law. In theory, the police could take up any activity that contributes to the objectives, not only the tasks that are mentioned in the law. Those activities include prevention work and cooperation with other actors, including the private sector (see the box below for a simple example).

Tasks versus objectives: the case of burglary

When the police is expected to execute tasks, they would focus on criminal investigations of burglars. After all, this is what the task 'enforcement of the law' included in the Police Act implies. However, when the police is expected to achieve objectives, the exact activities are not prescribed in the law. In this case, the objective could be to lower the rate of burglaries. It is up to the police to decide what activities best serve this objective. In practice, these turn out to be burglary prevention campaigns rather than criminal investigations (see Van der Vijver et al., 2001, p. 39).

2.3 Reasons for the regional police forces' level of discretion

The Secretary of the Interior needs to translate the tasks of the police into objectives for the regional police forces while accounting for the unique position of the Department. After all, many of the powers to manage the police - to set goals and priorities - are decentralized to the regional level (see Section 1.3). As we will argue below, this specific institutional structure is in the interest of the Department. By purposefully leaving considerable discretion to the regional level, the Department can make use of the expertise at that level in achieving the objectives for the police. We discuss three reasons why the forces' experience with policing under regional conditions is essential to effective police work.

Differing regional conditions leave room for regional priorities and methods

Crime and public order problems differ from region to region. The forces need detailed knowledge of local circumstances to be effective. The police needs to know who is who, who lives where, who deviates from the norm, and who had which role in what happened. They need to build up social networks and they need to be aware of cultural patterns in the neighbourhood (Van der Torre and Stol, 2000, Chapter 11). Clearly, the local forces have much better knowledge of the regional conditions. Therefore, it makes sense to give the forces some freedom in setting their own priorities and choosing their own methods.

The need to ration the supply of police services

There is structural excess demand for police services. Once taxes have been paid, the price for additional police services is often low for individuals or groups of people. Given limited police resources, there is an opportunity cost to reallocating resources from one activity to another

activity. However, these costs are often not evenly distributed. For instance, a storekeeper would rather see the police spending more time on reducing shoplifting than on enforcing traffic laws. Clearly, road workers have different interests.

The forces have better information to weigh competing calls for service than the Department. Their information advantage is based on their training and experience with police work in their specific locality. Therefore, it makes sense to leave the forces some flexibility in deciding whether a call is worth spending time on. This is not an easy job since the police will always dissatisfy some people.⁹

The nature of police work defies exact task descriptions

A policy of full enforcement of criminal law would imply that the police are required and expected to enforce all criminal statutes at all times against all offenders. It suggests that the police are without authority to ignore violations, to warn offenders when a violation has in fact occurred, or to do anything short of arresting the offender and placing a charge against him for the specific crime committed. There are a number of reasons why there is room in the law to tolerate criminal behaviour (Goldstein, 1963):

- *Ambiguity of the law.* Criminal law is often expressed in broad terms, which makes it difficult to render a clear interpretation of the legislature's intentions. Ambiguity may be intentional so as to provide greater flexibility in enforcement. It may also result from a failure to envisage the day-to-day problems encountered by the police or it may simply be a result of language limitations.
- *Enacting laws as a symbolic act.* Some laws either cannot effectively be enforced or are not meant to be fully enforced (think of the drug laws before they were changed). A lawmaker may enact such laws for several reasons. Maybe she does not want to be known as someone who does not support a law against reprehensible behaviour. Maybe she also does not want to be known as someone who repeals a prohibition because it is not really meant to be fully enforced. She derives her comfort from enactment of the law, not from enforcement of the law. In line with this behaviour, new legislation declaring a form of conduct to be criminal is rarely accompanied by an appropriation to support the resources for its enforcement. It falls to the police to live with the law without enforcing it.
- *Obsolescence of the law.* The law does not get updated every time reality changes. Therefore, there are obsolescent laws that the police cannot enforce or is not expected to enforce.

⁹ Where individuals or groups have conflicting expectations of the police, research conducted in the United Kingdom tends to confirm that improving satisfaction can be a 'zero-sum game' (FitzGerald et al., 2002, p. 113).

In theory, the Department could exactly prescribe the forces when and how to act. However, the unpredictable nature of the demand for police services defies exact task descriptions. The police face unexpected situations all the time. Clearly, there are patterns in criminal and disorderly behaviour, but the precise circumstances differ every time. A police officer is trained to handle these ever-changing situations and builds up expertise over time. Therefore, it makes sense for the principal to leave the police some flexibility.

To conclude, it is in the interest of the Department to actively involve the forces in the process of translating broadly defined objectives into workable goals and methods. Discretion requires the forces to think about what they want to achieve and the most effective way of reaching that goal. In other words, discretion comes with a responsibility for evaluation and innovation.

2.4 The dimensions of overall police performance

Given its unique relation to the regional police forces, what are the Department's objectives for regional police work? In this section, we will argue that the Department's focus is on the overall performance of the forces. The forces should provide efficient, effective, fair and decent policing within the framework of broadly defined outcomes of regional police work.

Efficient and effective policing

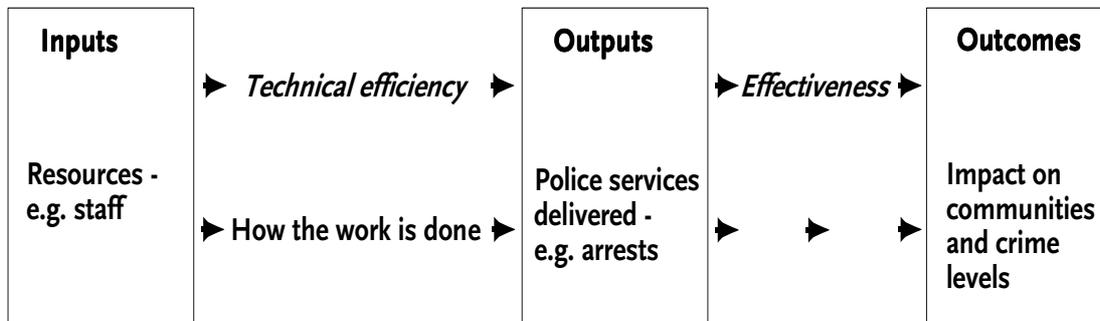
Because of the great degree of discretion at the regional level, the Department of the Interior has good reasons to define the forces' objectives in terms of desired effects of police work (outcomes like a change in the level of crime) rather than to provide detailed descriptions of desired outputs (e.g. number of fines) or inputs (e.g. time spent on the streets, 'blauw op straat'). An outcome is meant to reflect the effect of policing on society. In practice, outcomes also reflect the effects of many other actors and factors.

By focusing on the effects of police work, the Department relies on the forces' expertise to:

- *Choose those outputs that contribute most to the desired outcomes (maximum effectiveness).* Outputs include neighbourhood patrols, criminal investigations, prevention programs, etc. For instance, a force may choose to focus on prevention programs rather than repressive means to fight crime.
- *Organize inputs so as to maximize output (maximum technical efficiency).* For instance, a recent effort to increase technical efficiency refers to the frequency of coffee breaks, the level of reimbursement of travel expenses and the way the car fleet is purchased (TK 26345, 2002, p. 6). Another example is the choice to have one officer on a beat rather than two.

The Department's concern is with the outcome of these two simultaneous internal processes. In other words: efficiency and effectiveness in terms of desired effects of police work is one of the Department's main concerns (see Figure 2.1).

Figure 2.1 Efficiency, effectiveness and the relation between inputs, outputs and outcomes



Source: Based on PSP (2000), p. 11.

Fair and decent policing

The police is bound by the imperative of providing assistance and delivering justice in a fair and decent manner. The rules of fair and decent policing include issues such as limitations to the powers of investigations and the use of force and non-discriminatory policing. They also include more informal rules such as providing courtesy in all public contacts. The standards of conduct are constraints as they tend to slow down police work. Clearly, the police cannot always be friendly. However, even when delivering justice to hardened criminals, the forces operate under the imperative of fair and decent policing. Without such constraints, police officers would search houses without a search warrant, seek out criminals without reading them their rights, allowing them to call their attorneys, or releasing them in response to a writ of habeas corpus. The Gestapo was very efficient in maintaining public order and investigating crimes, however defined, but its work does not appeal to our idea of fair and decent policing.

The issue is not that an efficient police force would show no positive interaction with citizens whatsoever. To be able to do its work, it is in the interest of the police to maintain good relations with at least some citizens. These citizens may be helpful in providing information to the police, for example. Moreover, a highly effective police may also enjoy confidence among those who are not confronted with police brutality. The point is that the police should work indiscriminately with all citizens according to nationally defined standards. All citizens should get equal treatment under similar circumstances.

A definition of the objectives

Based on the above discussion and the Safety Program published by the (outgoing) Balkenende administration (Justitie and BZK, 2002), we deduce the following definition of the objectives:

The government's objective for the regional police forces is the reduction of crime, promotion of safety and reduction of disorder in an efficient, effective, fair and decent manner.

This definition of the objectives implies a focus on overall performance of the forces. Clearly, when assessing performance, the desired outcomes can be made more concrete (for example the desired contribution of the police to traffic safety and the quality of living in the neighbourhood).

The authorities at the regional level attach weights to the outcomes of regional police work; the Department's concern is whether regional police work meets the above-mentioned standards of policing. After all, the forces have knowledge of regional conditions and policing unmatched at the national level. Such a focus of the Department agrees with the status of the forces as independent bodies with their own administrative framework of mayors and public prosecutors. It also fits with the recent shift in the Department's policy away from prescribing policy and management priorities towards setting standards for regional police work (Terpstra, 2002, Chapter 2).¹⁰

2.5 Conclusions

As principal of the regional police forces, the Interior Department's objective is to make sure that they do their work in an efficient, effective, fair and decent manner. Their performance can be judged against outcomes deduced from the overarching objectives of the police: reducing crime, promoting safety and reducing disorder.

In line with this definition of the objectives, the Department relies on the forces' expertise to choose those outputs that contribute most to the desired outcomes (maximum effectiveness) and organize inputs so as to maximize output (maximum technical efficiency). The great degree of discretion of the forces is based on their superior expertise in deciding (1) whether to tolerate criminal behaviour or not, (2) how to handle criminal or disorderly behaviour, (3) how to handle competing calls for service (4) what priorities to set and methods to use in the region. In the next chapter, we will discuss the way in which the Department supports the forces' efforts to improve their performance.

¹⁰ Terpstra (2002, p. 57): "... the focus of management has shifted from national priorities to essential conditions for the organization." Translated from: "... de aandacht [is] verschoven van sturing op beleidsthema's naar sturing op organisatorische randvoorwaarden."

3 Realizing the objectives

3.1 Aim and structure

How does the Department of the Interior support the police forces in their efforts to improve their overall performance? In this chapter, we discuss the Department's policies to realize this objective. First, we introduce the challenge facing the Department in Section 3.2. The Department has good reasons to provide the forces with great discretion, but is responsible for their performance at the same time. We describe how the Department addresses this tension in its policies towards the forces in Section 3.3. We discuss other factors affecting the forces' motivation to work towards their objectives in Section 3.4. Section 3.5 concludes.

3.2 The Department's challenge

The Interior Department may be responsible for efficient, effective, fair and decent regional police work, the forces know 'where policing is all about' and 'what is really going on in the region'. Clearly, this is not an easy position to be in. Below, we argue that the Department will need to rely on incentives rather than rules to realize its objectives for policing, simply because it does not have the expertise to prescribe the forces how to organize their internal processes. By devising well-targeted incentives, the Department can support the 'intrinsic motivation' of the forces to improve their performance.

3.2.1 The rationale for providing incentives

The Department may set some rules to support improvements in performance such as compulsory outsourcing of the car fleet, but these rules have only limited reach. The desired discretion of the forces does not facilitate the establishment of such rules. After all, the Department does not have the relevant information to see what works best. Is patrolling with one officer rather than two a good idea to improve efficiency? Are prevention campaigns the best way to lower the rate of housebreaking? Without the expertise in policing and specific regional circumstances, the Department will have difficulty making these judgements.

If the Department is not able to prescribe how the forces should improve their performance, then it will have to rely on incentives to ensure that the forces achieve maximum value for money. Incentives do a call on the creativity of the forces. Given their information advantage, the police forces are in a much better position to identify opportunities to improve their performance (and to implement these ideas effectively). Incentives challenge the forces to find the best way of improving performance. As such, incentives can support the 'intrinsic motivation' of the forces.

3.2.2 The need to support 'intrinsic motivation'

Clearly, police employees are not only driven by external incentives in the form of financial rewards or career perspectives. People may choose to work for the police because of an inherent tendency to seek out challenges, to extend and exercise their capacities, to explore, and to learn. They are motivated to do their work well by factors that are intrinsic to them. Similarly, the police forces can be seen as 'intrinsically motivated' to realize their task in their specific region. If their motives are in line with the objectives for the forces, then the Department task of improving police performance through 'instrumental motivation' can be alleviated.

Indeed there are numerous examples of initiatives to improve the efficiency and effectiveness of police work. One major initiative is the peer-evaluation that is part of 'quality management' ('kwaliteitsstelsel politie'). Every four years, each force conducts a self-assessment. The self-assessment follows the 'quality model' of the Dutch police.¹¹ The self-assessment is critically reviewed by a team of auditors from the police and someone from outside the police. One year later, a review committee assesses what issues have been taken up after the self-assessment and audit. The committee consists of a mayor in command, a chief public prosecutor, a chief of police and someone from outside the police. The findings of the committee are sent to the Secretary of the Interior and the Police Inspection. The evaluations have resulted in some critical remarks about the organization of the police, including a lack of focus on results and low quality of management (FRP, 2000).

Such initiatives can be supported by recognizing and rewarding efforts to improve police performance. There are basically two reasons for performance evaluations and accompanying incentives:

- *Feedback.* Performance evaluations provide feedback to the forces on their perceived strengths and weaknesses. For example, forces may be unaware of whether their actions are satisfactory; an important purpose of performance evaluations is to tell the forces what areas they can improve in (Prendergast, 2002, p. 5117).
- *Recognition.* When the efforts that go into improving performance go unnoticed, the drive to identify and share best practices may be weakened. In a normal market there is an automatic reward for high performance in the form of higher returns to investment. With the monopoly

¹¹ The 'quality model' is an adapted version of the EFQM Excellence Model (as worked out by the Institute Dutch Quality, INK). The model describes the interaction between the determinants of quality and results. Areas include leadership, human resource management, strategy, resource management, processes, employee satisfaction, results, customer satisfaction and citizen satisfaction.

status of regional police forces there is no such reward.¹² Improving performance is a process of trial and error that comes with high costs. It takes considerable investment in data systems, entry of data, methods to analyse the data, and informal exchanges of experiences within and between forces. Based on the findings painful changes may need to be made in the organization. The accompanying costs can either be financial (e.g. investment in technologies for more efficient scheduling) or immaterial (e.g. stress stemming from effective personnel management). In the process, enterprising individuals within forces may have to overcome bureaucratic barriers and a continuous, strong pressure to respond to the issues of the day. Moreover, some benefits of the effort go to others ('external effects').

Therefore, incentives are instrumental in sustaining the motivation to improve police performance. That way, the Department can provide a stimulus - necessary to signal exceptional performance and to overcome the multiple barriers to action. Incentives do not replace but complement intrinsic motivation. The rewards can be either financial (e.g. performance rewards) or immaterial nature (e.g. status). As we will discuss later, it is essential that the incentives are well targeted. Poorly designed incentives may result in 'intrinsic frustration' rather than support intrinsic motivation (see the box on the Policing for London study in Chapter 5). However not providing incentives may not be appropriate response to the challenge of designing effective incentives. Indeed, there is some evidence from psychological research that a lack of incentives may have a negative effect on effort:

... because of systems that are insensitive to performance, some employees can vary their performance substantially with little effect on tangible reward. A low level of performance is often observed in such situations. Cognitive evaluation theory would explain this result as a lessening of intrinsic interest. Alternatively, employees may learn that their performance is largely irrelevant to reward, and such learned helplessness may have a detrimental effect on task performance. (Eisenberger and Cameron, 1996, p. 1164).

3.3 Incentives from the Department's policies

How does the Department address the challenge of taking responsibility for the performance of police forces that have a great degree of discretion? The Interior Department has two policy instruments: the distribution of budgets and the exercising of oversight. In this section, we

¹² The police forces may lose 'market share' from 'competitors' in the private security sector however. Indeed, the number of employees in private security per 10 police officers has risen from 2.7 in 1990 to 4.6 in 1999 (Source: CBS Statistical Yearbook, 2002). See Van der Vijver et al. (2001), p. 76 for a discussion of factors that drive this trend.

analyse whether these policy instruments indeed support the motivation to improve performance.

3.3.1 Distributing budgets

The distribution of budgets between the forces is dealt with separately from everything else. There is no direct link between a force's efforts to improve performance and budgeting decisions. Budgeting is rule-based. Budgets are adjusted based on a funding formula (described below). Therefore, the budget system provides no incentive to improve performance (BZK, 2001).¹³

The funding formula provides the number of budget-units ('budgetverdeelenheden' or BVEs) for each police force based on a number of regional characteristics. These characteristics include (1) miles of roads, (2) number of housing units, (3) total population, (4) number of non-Western immigrants, (5) population density, (6) frequency of residential moves, (7) number of shops. Frequently, the funding level for each budget-unit is adjusted. For 2003, a budget-unit is equal to EUR 61,258. In principle, each budget-unit is equal to 0.94 FTE, but the forces have some freedom in allocating their budgets. There are also some additional resources for forces that are faced with a region-specific workload, for example security of the harbor (Rotterdam-Rijnmond).

3.3.2 The Policy and management cycle

The Department's powers to oversee the performance of the regional police forces are embedded in the Policy and management cycle ('Beleids- en beheerscyclus politie', described in article 43a to 43d of the Amendment of the Police Act of October 12, 2000.). Below, we first describe the formal process, then we discuss the way the cycle works in practice. We conclude with a discussion of the incentives that result from the cycle.

The formal way the cycle works

In 1998, the first cycle started with publication of the 'Quadrennial Policy Plan' ('Beleidsplan Nederlandse Politie'), produced in cooperation with the Department of Justice (BZK and Justitie, 1998). In this Plan, the Secretary and the Justice Minister stated the most important outcomes of police work for the period 1999-2002. The outcomes included a decrease in youth crime and street violence, better help to victims of crime, and improved traffic safety. Other outcomes referred to supra-regional activities such as organized crime and international coordination. The Department restricts itself to setting priorities in general terms. It is to the forces to decide

¹³ Budgeting and oversight get more closely related in the situation that forces are under 'special oversight' of the Department ('preventief overzicht', applied to six forces in 2000). In that case, the Secretary of the Interior needs to agree with (changes in) the budget (In 't Veld et al., 2001, p. 55).

which outcomes should get less priority (although none of the forces gathers information about the backlog in any of their activities on a systematic basis, see Algemene Rekenkamer, 2003a, p. 23).

The Policy Plan provides the basis for the Annual Policy Circular ('Landelijke Politiebrieven') that describes the priorities of the forces for that year in more detail. The new Safety Plan includes a new set of desired outcomes of police work for the period 2003-2006 (Justitie and BZK, 2002). National priorities include (1) Cracking down on persistent offenders and youth criminality, (2) Strengthening criminal investigations and law enforcement, (3) Strengthening visible control in public space, and (4) Intensifying targeted prevention projects (Justitie and BZK, 2002).

The forces need to provide the Secretary with information that shows their efforts to realize the desired outcomes. The Secretary can prescribe explicit rules for the way the forces collect and provide these data. Currently, the Department prescribes the use of uniform performance indicators included in the Information Model Police ('Informatiemodel Nederlandse Politie', INP), a comprehensive effort to capture all police work in terms of results. The INP includes a mixture of 184 outcome, output, throughput and input indicators. Since 2001, the annual report of the police reports developments in some of the INP measures ('Jaarverslag Nederlandse Politie').

There are regular meetings between the Secretary of the Interior, the Justice Minister, the mayors in command and the chief public prosecutors to discuss progress in achieving the prioritized outcomes stated in the Policy Plan. To ensure execution of the Policy Plan, the Secretary of the Interior can give directions that need to be followed by the mayor in command ('aanwijzingen'). The mayor in command may need to change the police organization, the budget, the annual account or the police plan in accordance with these directions. When a force's performance still does not meet expectations, the Secretary of the Interior can fire the Chief of Police and (in the near future) the mayor in command (Justitie and BZK, 2002, p. 82).

The Secretary also has the authority to prescribe measures aimed at improving overall efficiency of the forces. The Quadrennial Policy Plan includes some suggestions (BZK and Justitie, 1998, Chapter 6 and 7). Suggested measures include a better method of scheduling work times, mandating one-person street patrols, better processing of crime allegations and more cooperation between forces. Another possible measure is the introduction of more flexible working hours (ibid., p. 38).

How the cycle works in practice

There is a wealth of data on the way the Department's Policy and management cycle works in practice. Three extensive studies have been conducted in this area: In 't Veld et al. (2001), Terpstra (2002) and, most recently, Algemene Rekenkamer (2003a). The conclusions of the studies are strikingly similar. The Department's oversight of regional police work is at best characterized as 'light-handed'. The collection of data needed to achieve greater accountability creates a tremendous bureaucratic burden, but few benefit in terms of meaningful information on performance. The forces' operational decisions are not affected by the Department's oversight.

The term 'national priorities' suggests that the Department does not only aim at improving the overall performance of the forces (in line with Chapter 2), but also wants to influence the relative weights attached to outcomes of regional police work. After all, the national priorities are not exclusively related to supra-regional police work such as fighting international crime (i.e. aimed at influencing the allocation of resources between regional and supra-regional police work). Moreover, the outcomes cover several but not all of the dimensions of regional police work.

A policy of influencing priorities would go against the logic of the decentralized system of regional police authorities in which mayors and public prosecutors are represented. Indeed, within the forces, people see a great discrepancy between the subset of desired outcomes that are coined 'national priorities' and the problems at the regional level (Terpstra, 2002, p. 52). However, in practice, there is no (credible) prioritization of outcomes of regional police work at the national level. Indeed, the national priorities in the Quadrennial Policy Plan were chosen so that no conflict would result with regional priorities (Algemene Rekenkamer, 2003a, p. 19). But there are many indications that there is no link between national priorities and regional circumstances:

- The original set of national priorities was not based on any analysis of the nature, volume and location of crime (Algemene Rekenkamer, 2003a., p. 12).
- There is a yearly adjustment in priorities, whereas it is hard to imagine that the safety situation changes so quickly that the relative weight of different outcomes need to be changed on an annual basis.
- Efforts of the forces on any of these national priorities do not lead to an adjustment in the national priorities (Terpstra, 2002, p. 55). The list of priorities tends to expand over the years: no priorities are dropped. The Annual Policy Circular mainly leads to an accumulation - and therefore depreciation - of 'national priorities'.

- The policy cycle does not even allow time for any meaningful coordination of priority setting for regional police work (In 't Veld et al., 2002, p. 13). The planning cycle of the forces does not match with the national policy cycle.

Therefore, the suggested prioritization of outcomes of regional police work at the national level is not realized. This finding could be seen as support for a focus on overall performance of the forces. After all, the focus of the Department is not on the relative weights attached to concrete goals set at the regional level. Indeed, the Department is moving away from interfering with regional policy plans towards finding a system of accountability with which to measure achieved results. The proposed performance contracts are in line with this change in strategy.

Perhaps because of the tension felt between the suggested prioritization at the national level and the discretion of the forces, the national priorities are not sufficiently specified to make them verifiable.¹⁴ Often the goals have no time horizon, no reference point, and cannot be quantified. Moreover, some of the terms used are ambiguous (for example, what is the definition of 'youth?') (Algemene Rekenkamer, 2003a, p. 12 and 13).

Therefore, meeting the national priorities has little practical implications beyond the paperwork that comes with it. The Department just conducts a formal check whether national priorities are included in the regional policy plans ('nota van bevinding'). The Department is not able to check actual compliance with the national priorities. Despite the abundance of data, the Department is not able to develop a clear picture of the forces' performance. This is partly the result of the fact that most data is incomplete and unreliable (BZK, 2001). Indeed, the paying out of the performance rewards has been delayed with a reference to data problems (see Chapter 6). However, a closer look at the way the Department processes these data reveals that even perfect data would not help:

- *There is no indicator of overall performance.* A force may do well according to one of the 184 INP indicators, but not so well according to another indicator, making it hard to draw conclusions.
- *'Value for money' is not being measured.* Reporting outcomes such as a change in the 'feelings of safety' across police regions is not informative for three reasons: (1) it implicitly assumes that all forces are equally affected by the many external actors and factors that co-determine such outcomes; (2) there is no relation with resources spent, whereas forces differ in the way they allocate their resources to alternative goals (Algemene Rekenkamer, 2003a, p. 47); therefore it is

¹⁴ In 't Veld et al. (2001), p. 93: "Most of the policy themes are put in such general and logical terms that each force pays attention to these priorities anyway". Translated from: "De meeste beleidsthema's zijn dermate globaal en logisch dat ieder korps sowieso aandacht besteed aan deze prioriteiten".

unclear what the forces achieve per euro or per officer; (3) it assumes that the police funding formula is correct in distributing resources among the forces, whereas some forces may be underfunded in comparison to other forces.

Consequently, the information collected is not used by either the Department or the forces. The three studies into the oversight of the Department suggest that the national priorities do not have a real effect on operational decisions of the forces.

Conclusion: lacking focus on ‘value for money’ creates little incentive

Without means to produce meaningful information on the performance of the forces, oversight cannot provide incentives to improve performance. It is unclear whether the forces are able to achieve the desired outcomes of regional police work in the most efficient, effective, fair and decent manner. Consequently, the efforts of the forces to improve their performance are not supported by the cycle. Instead, the inability to identify the results of such efforts may breed misunderstanding or even imputation between the Department and the forces. The box below provides some examples of the lack of meaningful interaction. The examples are based on interviews that we conducted for this study and newspaper articles.

Some statements illustrating how the Department’s information problem may breed a mutual lack of understanding

- “We cannot achieve those results with only 4,000 additional police officers”, see ‘Handvol extra bekeuringen de man’, *NRC Handelsblad*, December 31, 2002.
 - “We are very busy with our own regional problems, therefore you cannot expect from us to perform well on all the outcomes prioritized at the national level”, see in ‘t Veld et al. (2001, p. 10).
 - “Our region is different from other regions; we just look inefficient.”
 - “We are already changing our strategy and organization; you are too late to judge us fairly”, see Algemene Rekenkamer (2003a, p. 42).
 - “We have just started this new policy; you are too early to judge us fairly.”
 - “We are limited in achieving our targets because of problems elsewhere” (such as the limited capacity of the prosecution council and major events), see ‘Burger voelt zich onveiliger - onvrede over optreden van politie groeit’, *NRC Handelsblad*, June 21, 2002.
 - “We would like to follow your policy, but we are confronted with competing demands from other principals” (such as the mayor), see Terpstra (2002, p. 77).
 - “We put quality first, we just look inefficient.”
 - “Your information is outdated, we are dealing with a totally different environment now.”
 - “We are professionals; we know what we are doing.”
-

3.4 Other incentives

There may also be other external incentives that affect the behaviour of the forces (see the box at the end of this section for a story on a somewhat surprising type of financial incentive for the police in the United States). If these incentives work in the right direction and are sufficiently strong, they could partly offset the lack of incentives from the Department's policies.

A source of incentives is publicity. The media spend considerable time on criminal and disorderly behaviour and the response of the police. The status and future job perspectives of both the mayor in command and the chief of police partly depend on the type of media coverage they receive. Someone's image in the media may well be an input to tenure decisions made by principals such as the Department. Therefore, police administrators may be inclined to alter the force's course of behaviour if that results in good media coverage.

Lacking any better measures and information, performance of the police is mainly measured by rough indicators such as the crime rate, the crime clear-up rate and incidents ('police found drugs', 'man bleeds to death after assault', 'police officer suspended after use of violence', etc.). Prendergast (2001) provides evidence that strong media attention to incidents of police violence changed the behaviour of the Los Angeles Police Department (LAPD). The LAPD reduced crime-fighting activities in an attempt to avoid further negative press and investigations. The use of force per arrest fell, and so did the number of officer-involved shootings per officer. This trend towards 'cautious policing' coincided with an increase in gang-related homicides. Prendergast's analysis suggests that it was in the interest of the LAPD to alter its behaviour because of the strong media attention on incidents.

The Dutch police may be influenced by the media in a similar way. If media coverage is indeed a strong incentive, then we may expect to see a bias towards visible output that get lots of positive coverage (often called 'symbolic policies'). It may explain the campaigns of preventive body searches in Rotterdam and Amsterdam. These highly involving actions created a lot of media attention, whereas their impact on safety and crime relative to other ways of using policy capacity is unclear. But the anecdotal evidence on the behavioural response of the police is not clear-cut. The widespread use of the crime clear-up rate does not appear to lead to 'gaming' of the indicator by focussing on easy-to-solve crimes (for an analysis of the trends in this indicator see Wiebrens and Essers, 1999).

To the defence of the police, media attention may also frustrate their work through sudden changes in the goals and directions set by the regional police authorities or the Interior Department. *Ad hoc* policy priorities may reflect the issues that are fashionable that week or year

rather than a thorough analysis of crime problems. That seems a more plausible explanation of the frequently changing priorities than a change in the safety situation (see Section 3.3).

Incentives from media attention may explain the emphasis policymakers place on patrolling the streets ('blauw op straat') at a cost of not-so-visible crime investigations (an inefficient allocation of police resources according to some experts, see Boom, 2002). It may also explain the sudden allocation of police resources towards airline passengers who smuggle relatively small amounts of drugs from the Netherlands Antilles after extensive press coverage ('bolletjesslikkers affaire').

To conclude, another source of incentives, media attention, does not seem to stimulate efforts to improve performance either. In the absence of better measures of performance, it may lead to ad hoc policies. It may also lead to a bias towards visible police outputs to get positive coverage and to over-cautiously policing to avoid negative coverage. After all, letting hard core criminals go is likely to get less negative attention than a violent attempt to arrest them. However, we have to be cautious before drawing hard conclusions since there is no systematic empirical evidence on the nature of the influence of the media on police behaviour in the Dutch context.

An additional preliminary conclusion we can draw is that making a set of more consistent performance indicators publicly available may improve the forces' incentives to improve their performance. We will get back to this issue in Chapter 10.

Do financial incentives affect police behaviour? Asset seizure and forfeiture in the United States

In the United States, the federal government introduced direct financial incentives for local police agencies to 'buy' police time for a national priority: the 'war on drugs'. The Comprehensive Crime Control Act of 1984 allows local law enforcement agencies to keep a portion of the net proceeds of forfeitures they help make under federal law. These forfeitures are mainly meant to take away the proceeds of drug dealing. For example, the police frequently confiscates cash, cars and other valuables in drug arrests. Previously, forfeited assets had been handed over to the federal government in their entirety (the Dutch equivalent is the 'Pluk ze' operation).

There are indications that these direct financial incentives indeed led to a strong behavioural response of the police. Immediately following passage of the Act, federal forfeitures increased dramatically. The amount of revenue deposited into the Department of Justice Assets Forfeiture Fund, for example, soared from \$27 million in 1985 to \$644 million in 1991 - a more than twenty-fold increase. Since 1986, the Department of Justice has distributed over \$2 billion in cash and property to local police agencies. Additional revenue comes from forfeitures done under state law, which adds to the total intake. Some police departments single-handedly took in several million dollars for their own use. There are concerns that the strong financial incentives lead to overzealous law enforcement harming innocent citizens and of displacement of state and local police priorities.

Some of the growth in forfeitures may be due to a simultaneous policy change, however. Although there are strong indications that the financial incentives mattered, it is unclear what their exact effect is in relation to the intensification of the 'war on drugs'. An empirical study separating effects of the broader policy change and the financial incentives would be necessary to turn the above indications into hard evidence.

Sources: Dunn (2000), Skolnik (2001).

3.5 Conclusions

Although the forces take initiatives to improve their performance, recognizing and rewarding such efforts may support their motivation to do so. External incentives can provide useful feedback on successful initiatives and compensate for the costs of improving performance. Neither the Department's policies nor other sources of incentives such as publicity appear to support the forces' intrinsic motivation, however. The lack of incentives may negatively affect the drive to improve performance. Therefore, there seems to be a need for change in the Department's policies towards the forces. In the next chapter, we will discuss whether the available empirical evidence supports this assertion.

4 Empirical evidence on police performance

4.1 Aim and structure

In Chapter 3, we concluded that there seems to be a lack of incentives to support performance improvements. If the lack of incentives makes a difference, we should be able to notice that police performance stays behind. In this chapter, we will review some of the empirical evidence. As we will see, empirical analyses into the performance of the forces are scarce. The lack of evidence itself indicates that improving police performance has not been high on the agenda.

4.2 Hypotheses

We will assess whether the following hypothesis can be supported:

Hypothesis 1. The regional police forces are operating at lower than maximum performance compared to foreign police operating under similar conditions.

Hypothesis 2. There are great differences in performance between individual forces.

Hypothesis 3. The forces do not show much improvement in their performance.

As a clarification of the second hypothesis: lacking effective oversight from above, there is no strong incentive to exchange best practices among the forces. In other words, the regional forces operate as 'islands' when it comes to learning how to improve performance. If this is the case, we expect to see great differences in performance between forces.

4.3 Crime clear-up rate Nordrhein-Westfalen vs. the Netherlands

Tak and Fiselier (2002) compared the crime clear-up rate between the German state of Nordrhein-Westfalen and the Netherlands. Their findings can help us to test the first hypothesis: are foreign police performing better than Dutch police?

Nordrhein-Westfalen is comparable to the Netherlands on many dimensions, including economic, demographic and social variables that explain crime (ibid., Chapter 1). The two regions have a similar level of recorded crime. Therefore, it is interesting to compare police performance between both regions. A comparison learns that the Dutch police performance seems to be much lower than that of the police of Nordrhein-Westfalen. In 1999, the German crime clear-up rate was 50.1 percent, whereas the Dutch rate was 13.4 percent (ibid., p. 26). Whereas the German clear-up rate has been increasing over the last 20 years, the Dutch clear-up rate has been on the decrease over the same period.

However, there are several possible alternative hypotheses that have to be tested before we can draw hard conclusions. First, there could be a difference in counting the number of solved crimes. The authors claim that such differences do not drive the results: the German and Dutch police use a similar definition of a cleared crime (ibid., p. 38). A second hypothesis is that German crimes are for some reason easier to solve than Dutch crimes.¹⁵ This hypothesis can be easily refuted: Dutch clear-up rates are consistently lower than German rates across many types of crimes (ibid., Table 5, p. 27). Third, it could be that the German police allocates many more resources to criminal investigations than the Dutch police (the number of officers per population is equal in both regions). We know that the Dutch police spends some 5 percent of their time on criminal investigations (OM, 2002). However, we do not know how the German police compares to the Dutch police in terms of allocation of time. We do know that research has shown that the amount of time spent on criminal investigations tends to be relatively low in many countries (Van der Vijver et al., 2001, p. 50). However, lacking exact data, we cannot explicitly test the hypothesis that the regional police forces are operating at lower than maximum performance compared to foreign police operating under similar conditions.

To conclude, a comparison of crime clear-up rates indicates that the Dutch police forces are relatively inefficient crime solvers compared to the police in the German state of Nordrhein-Westfalen. Somehow, the German police is able to produce more output than the Dutch police with a comparable budget and within a similar operating environment.¹⁶ As stated, the difference in performance could be (partly) due to differences in time spent on criminal investigations.

4.4 Weighted solved crimes per officer across forces

As an alternative to the crime clear-up rate, Wiebrens (2002b) presents the ‘weighted solved crimes per officer’. He weighs each crime that a force forwards to the prosecution council with the corresponding penalty in terms of prison days,¹⁷ adds up these ‘prison day equivalents’ and divides the resulting total by the number of police officers. The average penalty for a crime is a

¹⁵ There can be two reasons why German crimes are easier to solve than Dutch crimes: (a) compared to the German police the Dutch police may focus on hard-to-solve crimes.; (b) a difference in the ‘supply’ of crimes: either reporting behaviour differs (Germans report more easy-to-solve crimes) or there is a difference in the underlying supply of crimes (there are more easy-to-solve crimes in Germany).

¹⁶ Based on these data, we cannot directly test the claim that stronger incentives will result in higher efficiency. After all, Tak and Fiselier (2002) do not provide us with an analysis of the incentive structure within the police of Nordrhein-Westfalen.

¹⁷ Not all penalties take the form of incarceration. Other penalties include fines and duties. With the help of guidelines of the prosecution council each penalty can be converted into an equivalent number of prison days.

measure of the severity of a crime. Wiebrens's measure provides an indication of a force's efficiency in investigating crimes. His findings can help us to test the second and third hypothesis: are there great differences in performance between forces? and: do the forces show much improvement in their performance?

Data over 2001 show wide differences in performance between the forces. On average, officers in Noord Holland-Noord each produced 206 prison day equivalents compared to 315 days per officer in Flevoland.¹⁸ Not all of the difference may be attributable to efficiency however. First, large forces may enjoy scale economies in crime investigations compared to small forces. Wiebrens claims that the data does not support this hypothesis (*ibid.*, p. 902). Second, some forces may spend more time on crime investigations than other forces. The Algemene Rekenkamer (2003a, p. 47) provides data on allocation of time by four forces in different parts of the country over 2001. Based on these data, we conclude that there are no major differences between forces in time spent on crime investigations. Therefore, the differences in performance do not seem to be the result of different allocations of police time.¹⁹

A comparison of the performance over the period 1996-2001 gives a bleak picture of the development of efficiency of the forces. Over this period, the police budget grew by 17 percent. However, the number of prison day equivalents per officer decreased by 12 percent. The decrease cannot be explained by lighter penalties: the average penalties for crimes remained constant over this period. There are widely differing developments in efficiency between forces. One force improved its performance (Groningen), three forces remained constant (Amsterdam-Amstelland, Brabant-Noord and Gelderland-Zuid), the other 21 forces' performance worsened - with some forces showing a decrease of more than 25 percent (IJsselland, Gelderland-Midden and Noord Holland-Noord). It is unlikely that the forces differed greatly in changes in allocation of time to tasks producing 'prison day equivalents', making it more likely that there are wide differences in efficiency between forces.

To conclude, Wiebrens's analysis provides indications that the efficiency of most forces in crime investigations decreased substantially over the period 1996-2001. Some of the decrease may be due to changes in the allocation of time away from criminal investigations. A firmer conclusion

¹⁸ We assume that there are equal returns (in terms of prison day equivalents) to investigative efforts for different crimes. Under this assumption, it does not matter whether a force focuses on crimes that are relatively easy to solve. This force will solve more crimes, but each crime gets a lower weight, making the total prison day equivalents per officer comparable to a force that focuses on hard-to-solve crimes.

¹⁹ The absence of major differences in time spent on criminal investigations also refutes the idea that imperfections in the funding formula drives the apparent differences in performance. After all, the approximately equal share of resources allocated to investigations puts the forces on an equal footing.

is that there are wide differences in (changes in) efficiency between the forces. Moreover, some poorly performing forces were not able to catch up with other forces.

4.5 Overall productivity of the police, 1980-2000

Kuhry and Van der Torre (2001) are unique in deriving a measure for changes in overall productivity of the police. Productivity is defined as the efficiency in producing outputs (e.g. clearing up a crime). This measure matches with the concept of technical efficiency (see Chapter 2). It is measured as 'total factor productivity': the change in output that cannot be attributed to growth in labour (i.e. change in the total number of worked hours) or a rise in the amount of capital per unit of labour (e.g. a higher number of police cars per officer). Their findings can help us to test the third hypothesis: do the forces show improvement in their performance?

The authors derive the productivity measure in four steps. First, they develop an index for changes in the volume of output over time. Police products are multiplied with weights that reflect their relative use of police time. The size of the population is used as a proxy for those activities that do not produce easily measurable outputs (community policing, emergency assistance and some prevention activities). Second, the output index is multiplied with costs per unit of output (deflated with a general price index) to derive changes in the real costs per unit of output. Third, the resulting change in real costs per unit of output are decomposed in three parts: (a) change in real labour costs per FTE (b) change in real non-labour costs per unit of output, and (c) change in labour productivity. Finally, the difference between the changes in labour productivity and in non-labour resources per unit of output weighted with their respective costs shares gives us the rate of technical change, i.e. the change in 'total factor productivity'.

Figures 4.1 and 4.2 show the growth in labour and output and the change in labour productivity, technical change over the last 20 years. The figures exclude the effect of the introduction of automated detection and processing of traffic offences ('Wet Mulder') on productivity.²⁰ Clearly, labour productivity has gone down since about 1990. Somehow, the growth in the number of officers did not translate in a similar growth in the volume of output. The decrease in labour productivity was cushioned by capital deepening investments (the change in labour productivity exceeds technical change in Figure 4.2). Apparently, police work was positively affected by a higher amount of capital per unit of labour (such as police cars and information systems).

²⁰ We asked the authors to exclude the 'Wet Mulder' in their computations as the somewhat outdated weights used to derive the volume of output were made before the introduction of speed and red light cameras. The current weights are based on PKP (1988) and will be updated in the near future.

Figure 4.1 Changes in volume of output and labour, Dutch police, 1980-2000 (1990=100)

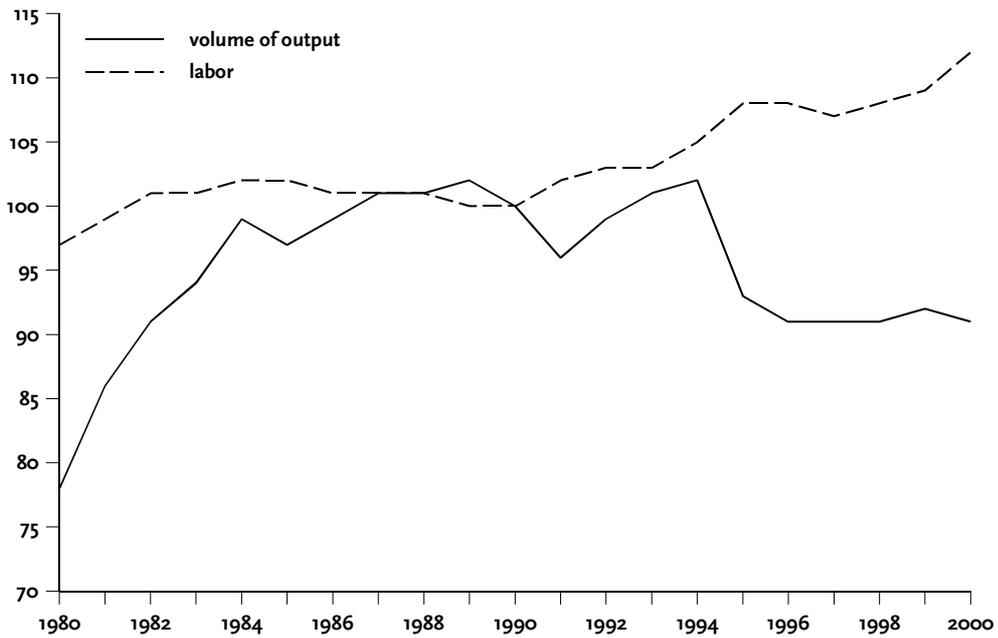


Figure 4.2 Labour productivity, technical change, Dutch police, 1980-2000 (1990=100)



Source: Based on Kuhry and Van der Torre (2001).

Again, the empirical evidence points at worsening police performance. The rate at which performance is decreasing is improbably high, however. According to this analysis, labour productivity has decreased by more than 15 percent over the period 1990-2000. What is going on here? Clearly, changes in the level of effort cannot fully explain this dramatic change. In the next section, we will discuss a number of policies that could have contributed to the decline in measured performance. Additionally, there may be some technical reasons for a downward bias in the productivity measures. First, improvements in the quality of (measured) output may not have been accounted for. Second, efficiency improvements may have been excluded because of the use of outdated weights underlying the index for the volume of output. For example, the introduction of computers may have made administrative police work more efficient.

4.6 Possible explanations for the decline in efficiency

All three studies provide indications that the efficiency of the police has been *decreasing*. Dutch crime clear-up rates have been declining over the last 20 years, whereas they have been stable or improving in other countries. The number of weighted solved crimes per officer has decreased over the last five years. Similarly, total factor productivity of the police has been on the decrease since 1990.

A priori, there is no clear explanation for the decline in efficiency. Given a lack of incentives, performance could stay behind, but performance should not necessarily decline - certainly not at this rate. Based on interviews, we found a number of possible explanations for the decline in efficiency:

- *Restraint on police output by the prosecution council.* Over the last five to six years, the prosecution council has been actively limiting the number of charges forwarded to the prosecutor because of a lack of capacity (for the specifics of this policy see the front page of *de Volkskrant*, September 27, 2002). An indication supporting this hypothesis is the decrease in the number of (non-technical) dismissed charges (from 23,366 in 1995 to 11,373 in 1999, see Huls et al., 2001, p. 339). After all, a decrease in the excess supply of charges would lead to a lower number of dismissals.
- *Greater administrative burden.* The administrative burden related to handling crime are said to have increased with greater attention to care for victims of crime (including 'Wet en Richtlijn Terwee' introduced in 1995) and to rules for criminal investigations (including 'Wet Bijzondere Opsporingsbevoegdheden', BOB, introduced in 2000).
- *Greater attention to labour-intensive fighting of organized crime.* Especially during the period that Hirsch Ballin was Minister of Justice (1989-1994) more attention was given to fighting organized crime. Resources were shifted from regional police work to the supra-regional and

national level. Often a single case involved many detectives. Consequently, output per officer declined.

- *Shift to community policing.* Another explanation for the decline in efficiency is a shift to activities with lower measurability of outputs such as community policing ('blauw op straat'). The above evidence is biased towards efficiency in the area of criminal investigations for reasons of measurability of output.²¹ Increased coordination with other parties in the neighbourhood such as youth work became more time consuming. Combined with a policy of highly tolerant policing on the street ('gedogen'), a policy officer shifted to this area will not produce much measurable output. This trend may have been strengthened by the shift from active work to back office functions resulting from the reorganization of 1993/94 (Rovers, 1999, p. 14).

Lacking further empirical evidence, we do not know which explanation is valid. If we would be able to directly test these hypotheses, then the key question would be whether these developments are desirable. If the administrative burdens indeed serve a purpose and more officers were shifted to community policing, how did that work out in terms of (feelings of) safety, for example? What happened to the effectivity of police work? Unfortunately, there are very few studies available on the effectiveness of police work. To the best of our knowledge, there is no study that looks into value for money in terms of effects of police work on society. This is not surprising given the difficulties related to measuring the effects of police work, even on things as crime (see McCrary, 2002 and Travis and Waul, 2002). Clearly, there is a need for empirical studies into the effects of police work rather than efficiency of a small set of tasks that produce easily measurable outputs.

4.7 Conclusions

Based on the scarce empirical evidence on police performance, we can draw the following conclusions:

- *There are strong indications of great differences in (changes in) efficiency between forces.* Moreover, some poorly performing forces are staying behind: they are not able to catch up with the other forces.
- *Either the Dutch police are working highly inefficient and are increasingly becoming more so - also compared to foreign police working under similar conditions - or efficiency just seems to be going down as a result of (a) capacity problems in the criminal justice system; (b) greater administrative burdens;*

²¹ Zoomer (2001, p. 23) describes the tasks of a community beat officer as follows: '[next to traditional police tasks, the officer spends time on] conferring with external partners (including the local administration, assistance, people in the neighborhood, proprietors and others), building and maintaining networks and stimulating the self-reliance capacity of citizens'. Clearly, none of these activities create much measurable output.

(c) greater attention to labour-intensive fighting of organized crime; and (d) a shift to activities with lower measurability of output such as community policing.

The first conclusion supports the idea that lacking effective oversight from above, the forces have little incentive to share best practices. Apparently, the forces operate as 'islands' rather than 'partners in crime' who share best practices. Lacking better information on the factors driving the 'productivity puzzle', we cannot be more specific in our conclusion about overall efficiency. Clearly, more empirical research in this area is needed. As we stated at the beginning of the chapter, the lack of empirical studies into police performance indicates that improving police performance has not been high on the agenda.

In Chapter 5, we will discuss how performance contracts between the Department and the forces could strengthen incentives aimed at improving performance.

5 The rationale for performance contracts

5.1 Aim and structure

The findings in Chapter 3 and 4 support the government's wish to improve accountability and incentives for the police. Although the empirical evidence is scarce, it all points at poor and decreasing police performance and great differences in performance between forces. Given the apparent need for change, is it logical to use performance contracts as an instrument to achieve the Department's objectives for the forces? In this chapter, we define conditions under which performance contracts could indeed be effective. We also review some of the scarce empirical evidence on the potential of performance contracts in the case of policing. Before launching into the discussion, we start with a definition of performance contracts.

5.2 What is a performance contract?

Performance contracts can be defined as follows:

In a performance contract, the principal specifies rewards and possibly penalties conditional on the extent to which the contracting organization meets targets on specific performance indicators.

It is important to note that the performance contracts are aimed at the organizational level. As such, they are different from contracts between the principal and contracting managers, so called 'managerial contracts'.

Despite their suggestive name, public sector performance contracts are rarely legally binding. In most cases performance contracting is implemented by administrative and managerial discretion rather than under a statutory or legal basis. They are mutually negotiated, or even implied, agreements between agencies to clarify undertakings of mutual interest (OECD, 1999, p. 14).

Examples of performance contracts

There are a number of performance contracts in place within the Dutch public sector. One of the most talked-about examples is the contract between Netherlands Railways (Nederlandse Spoorwegen, NS) and the Department of Transportation. The NS provides railway passenger services throughout the Netherlands. The contract was initiated in 2000. It included penalties for not meeting targets for frequency of services and punctuality. Indeed, the NS has been fined when targets were not met. In the latest version of the contract, these financial incentives have

been dropped, however ('Derde Verlengingovereenkomst bij het verlengde Overgangscontract II').

A second example - that excludes rewards or penalties - is the contract between the Informatie Beheer Groep (IBG) and the Department of Education. The parties entered into the first contract in 2002. The IBG is tasked with administering student loans and grants and the public transport student pass ('OV-studentenkaart'), providing information to students and a number of other, related activities. Performance targets are related to response time to phone calls and to objections in writing, customer satisfaction and the number of objections lodged by customers.

The two above examples are closest to the definition at the beginning of this section. There are several other, more implicit performance contracts for independent bodies and other relatively independently operating government agencies. 'Watered down' versions of performance contracts can be found in social security (e.g. between the Social Benefit Administration (UWV) and the Department of Labour, see Koning and Delen, 2003) and in the area of publicly funded research (e.g. between the institute for applied research TNO and the government, see Cornet and Van de Ven, 2003). Recently, there has been calls for the application of performance contracts in many other areas including the prosecution council, the provinces and municipalities.

5.3 Conditions for effective performance contracts

Performance contracts fit in with the 'intrinsic motivation' of the forces to improve their performance themselves and acknowledge the information asymmetry between the Department and the forces on the way to improve performance. Performance contracts combine the two policy instruments of the Department - distributing budgets and exercising oversight of the performance of the forces.

However, there are a number of non-trivial conditions that need to be met to turn performance contracts into an effective policy instrument: commitment of the Department and conditions for contract design. As we will discuss below, these conditions pose major challenges to the Department. If performance contracts can indeed be made effective, then their benefits need to be weighed against the transaction costs related to negotiating agreements and monitoring compliance.

Department's commitment to terms of contract and non-intervention

Commitment of the Department to the terms of contract and to a policy of non-intervention is necessary to create incentives:

- *Commitment to the terms of contract.* If the Department changes the terms of contract frequently, then the forces may start to anticipate on these changes. When the Department loses its reputation as trustworthy contracting partner, the performance contracts are not likely to strengthen incentives. Therefore, the Department has an interest in committing to the terms of contract, i.e. the way performance is measured and rewarded.
- *Commitment to non-intervention.* Withholding from intervening in the internal optimization processes of the forces during the contract period strengthens incentives. After all, the whole idea of the performance contracts is that the Department leaves the forces sufficient room for their own optimization processes. By following a policy of non-intervention the Department does not become a 'co-producer' who shares responsibility for the results. Thus the credibility of the Department as contracting party - and therefore the power of incentives for the forces - depends on the commitment to non-intervention.

The wish for flexibility in steering the police and adapting the terms of contract may go against the desire for commitment. A policy of non-intervention conflicts with the current habit of authorities to place sudden, new demands on the police - for example in response to media reports (see Section 3.4). As argued in Chapter 2, it is actually in the interest of the Department not to intervene in internal processes of the forces (except from rules for fair and decent policing). In principle, the Department's objective for the forces should provide for a stable set of benchmarks to be included in the performance contracts. However, it is unclear whether there is political will to commit to non-intervention.

Additionally, the Department may be tempted to change the terms of contract when there are bad experiences with the original contract design. Given the current state of thought on this topic in the Netherlands, it is imaginable that new performance indicators will be developed that better reflect the results of regional police work. By 'freezing' current methods in a performance contract, the Department will not be able to make use of these advances. Therefore the Department is faced with a tradeoff between commitment and flexibility. Committing to the terms of contract has the advantage of strengthening incentives, but goes at a cost of the flexibility to make use of future advances in measuring and rewarding police performance.

Contract design results in well-targeted incentives

Additionally, there are a number of conditions for contract design that need to be met if performance contracts are to be effective:

(1) *Performance assessment should strengthen incentives without inviting the forces to 'game' the evaluation:*

- The performance measures should pick up the results of effort in all of the relevant dimensions of police work so as to prevent biases in policing.²²
- Manipulation of data should not substitute for real efforts to achieve the performance targets (manipulation is not inconceivable given the recent fraud in higher education and employment placement services).
- The forces should have control over the performance indicators. There is no incentive without a link between effort and measured performance.

(2) *Measurement error should not weaken incentives by distorting the link between effort and performance.*

(3) *Targets and rewards should strengthen incentives without creating undesirable side effects:*

- Unrealistic targets should not undermine incentives.
- Bringing together several performance measures to determine whether to award the financial reward or not should not invite the forces to game the assessment (by only focusing on the easy targets, for example).
- Strong financial incentives do not put poor performers into a negative spiral. With strong budgetary consequences of poor police performance, citizens would pay twice for poor management within the police. After all, citizens do not have a choice among different police departments.

The importance of contract design is underlined by empirical research into this topic. In a review of almost 100 psychological studies into the effects of reward on motivation, Eisenberger and Cameron (1996) find that 'detrimental effects of reward occur under highly restricted, easily avoidable conditions'. Similarly, in an empirical study on the effects of performance contracts on productivity in Chinese state enterprises, Shirley and Xu (2001) find that the design of performance contracts makes all the difference. Successful performance contracts can improve productivity when they provide high-powered incentives, use sensible targets and signal commitment through features as long terms. Favourable environmental conditions can further strengthen thoughtful design of incentives. The more effective performance contracts are more likely in smaller, relatively well-running, competitive firms. In the absence of these features, performance contracts can hurt productivity. Thus, contract design can make or break performance contracts.

²² Assuming that substitution between tasks is possible; if tasks are complementary, an incomplete contract will not create a bias in police work, see Section 7.3.1.

A priori, it is hard to unambiguously answer the question whether the conditions for effective contract design can be met in the case of policing. We will have to rely on evidence in other countries. We review some of the existing evidence in the next section.

The transaction costs of performance contracts

Performance contracts are not costless (neither are other management regimes). If performance contracts can indeed be made effective, then the marginal benefits need to be weighed against the additional costs. The costs of performance contracts include time and other resources spent on negotiating various agreements and monitoring compliance, preparing periodic reports and statements, and maintaining contracting and monitoring staffs. In a study of performance contracts in New Zealand, Schick (1996, p. 24, 25) finds that a non-trivial part of the efficiency gains resulting from performance contracts may have been absorbed by high transaction costs associated with negotiating agreements and monitoring compliance.

5.4 Empirical evidence on the potential of performance contracts for policing

To the best of our knowledge, performance contracts for the police - including rewards and/or penalties - have not been applied anywhere around the world. Therefore, there is no direct empirical evidence on the effects of performance contracts in the area of policing.

The only study providing empirical evidence on the behavioural response to incentives we are aware of is the 'Policing for London' study (FitzGerald et al., 2002). This study is based on focus groups and in depth interviews with a large number of London police officers and managers. The study provides some interesting clues as to how well the 'contracts' worked for them. We provide a summary of the findings in the box below. The response from the London police shows that the performance measures could not simply be ignored. The contract introduced incentives to align behaviour with the performance measures. The authors do not provide information on any unexpected behavioural responses to the measures. They focus on the perception of the measures by the police instead. Most important conclusion is that performance measures result in frustration if they do not match with the locally perceived ideas about the priorities in police work.

Policing for London: behavioural response to performance targets with implicit rewards and penalties

Police officers and managers working with the Metropolitan Police in London were interviewed to see how they appreciated the use of quantitative performance measures. In 1999-2000, they had to comply with 9 Key Objectives set by the Home Secretary, 9 Metropolitan Police Charter Targets, 38 Best Value Performance Indicators from the Audit Commission, and 29 performance indicators used by the Police Inspection (HMIC). The interviewees pointed out that the focus on these performance indicators was no unqualified success. The measures forced the police into a crime fighting strait jacket and led to conflicts between national demands and local needs. These undesirable side-effects resulted in a loss of staff morale.

Implicit bias towards crime fighting

The performance measures generally gave primacy to crime-fighting objectives at the expense of order maintenance. Managers felt that the measures implicitly assumed a purpose of policing that did not match with their own experiences. There was a lot of pressure to meet targets on a narrow range of crimes, which only represented a minority of the calls they had to answer.

Tension between national demands and local needs

The performance measures imposed the same priorities across widely differing areas, regardless of variations in crime problems. This limited the capacity of local police to prioritize according to local need. Middle managers found themselves squeezed between centrally set priorities, on the one hand, and local needs and locally expressed priorities on the other. They felt the performance measures by which the central government judged them were not the criteria by which the public judged them.

Loss of staff morale

The crime-reduction targets which local managers had to deliver did not match with the workforce's understanding of what the job was actually about. Furthermore, the targets had little link with reality. They had been arbitrarily set in the absence of any firm knowledge about the means by which they could be achieved. The mismatch between centrally set performance targets and the workforce's sense of purpose - combined with the administrative burden related to reporting achievements - compounded the cynicism about management, and resulted in a loss of staff morale.

Source: FitzGerald et al. (2002).

5.5 Conclusions

In theory, performance contracts could leverage the Department's oversight of the forces' performance by tying budgetary consequences to improvements in efficient, effective, fair and decent policing. They fit in with the 'intrinsic motivation' of the forces to improve their performance themselves and acknowledge the information asymmetry between the Department and the forces on the way to improve performance. The resulting incentives can complement the 'intrinsic motivation' of the forces by providing feedback and compensating for the many

barriers that need to be overcome to improve performance. As such, performance contracts could foresee in the apparent lack of incentives for the forces.

The potential beneficial effect of performance contracts on police performance is conditional on the Department's commitment to the terms of contract and a policy of non-intervention and on the possibility of creating well-targeted incentives with the use of performance measures, targets and rewards.

The wish for flexibility in steering the police and adapting the terms of contract may undermine the power of incentives. A policy of non-intervention conflicts with the current habit of authorities to place sudden, new demands on the police - for example in response to media reports. Additionally, the Department may be tempted to change the terms of contract when there are bad experiences with the original contract design.

Designing a contract in a way that results in well-targeted incentives for the police may be even more challenging. There are many ways in which well-intended efforts to strengthen incentives could have undesirable side effects. Little is known about the behavioural effects of performance measures, targets and rewards in the area of policing, however. The current empirical evidence provides little guidance. Experiences in London show that a mismatch between performance measures and locally perceived priorities in police work may frustrate rather than support intrinsic motivation.

Moreover, the uncertain benefits of performance contracts will need to be weighed against the possible increase in transaction costs. Experiences in New Zealand show that high transaction costs are associated with negotiating agreements and monitoring compliance.

Step II. Are the proposed performance contracts likely to be effective?

In step I, we concluded that, in theory, performance contracts could be instrumental in achieving the Department's objectives for the forces - conditional on contract design, commitment of the Department, and transaction costs. We now focus on effective contract design. Are the proposed performance contracts likely to introduce well-targeted incentives?

In Chapter 6, we provide a description of what is known at this time about the design of the proposed performance contracts. In Chapter 7, we discuss the difficulties of evaluating police performance. In Chapter 8, we assess the likelihood that the performance contracts will be effective. The conclusions of the second step provide input for the third step in the analysis: what is the value of performance contracts for the police?

6 The proposed performance contracts

6.1 Aim and structure

In this chapter, we describe the proposed performance contracts between the Department of the Interior and the regional police forces. Section 6.2 provides a brief history of the current proposals and the plans for the near future. We discuss the design of the performance contracts in Section 6.3. Section 6.4 concludes.

6.2 Brief history and plans for the future

The performance contracts are meant to improve accountability and to create incentives at the regional level for the realization of the 2003 national police covenant. The covenant is meant to support the objectives set out in the 'Safety Plan' ('Veiligheidsprogramma') of the (outgoing) Balkenende administration. The overall aim of the Safety Plan is that 'there is indication that a decrease in crime and disorderly behaviour of about 20 to 25 percent is in sight by 2006'.²³ The covenant was signed on February 15, 2003 by the Secretary of the Interior, the Justice Minister and the mayors who are in command of the regional police forces ('korsbeheerders').²⁴ The parties agreed to enter into regional performance contracts by no later than July 1, 2003. The performance contracts may actually also include goals that go beyond the national covenant.

The main precursor of the performance contracts is the system of performance rewards that never got off the ground. In the national police covenant 1999, the Secretary of the Interior and the management council of the police ('Korpsbeheerdersberaad') agreed to introduce performance rewards by no later than 2001. In December 2000, the Secretary of the Interior asked the management council to advise on the design of performance rewards. In expectation of the advice, the Secretary stated that performance rewards would indeed be introduced in 2001 (BZK, 2001). The council presented its advice to the Secretary in June 2001 (KBB, 2001). Some two months later, the Secretary sends out a letter to the Parliament saying that he will adopt the proposed system (TK 26345). The first round of performance rewards would be paid in 2002. However, in November 2002, the Parliament was informed that payments of the performance

²³ Translated from Justitie and BZK (2002): "Met betrekking tot de objectieve veiligheid is in het programma aangegeven dat een vermindering van de criminaliteit en overlast in de publieke ruimte met - indicatief - circa 20 tot 25% vanaf 2006 in het vizier moet komen."

²⁴ The mayors were little enthusiastic when signing the covenant (*de Volkskrant*, February 15, 2003). "We should not pretend that the Netherlands is going to be safer with this kind of documents", according to mayor Weterings of Beverwijk. Brouwer, mayor of Utrecht, signed the covenant since, and we quote, "I cannot explain not signing. After all I go for more safety."

reward were put off to 2003 (BZK, 2002d). The Secretary said that incomparability of data did not allow for legitimate payments.²⁵ In the beginning of 2003, the Secretary of the Interior and the Justice Minister announced that the system of performance rewards should be changed given the new national police covenant (BZK and Justitie, 2003). The performance rewards will be integrated with the new performance contracts, which are to be signed before July, 2003. Until that time, the conditions for receiving the performance rewards are unclear.

6.3 The design of the performance contracts

The performance measures and targets in the performance contracts will reflect agreements made in the national police covenant (BZK and Justitie, 2003). The financial incentives for the forces will be largely based on the existing system of performance rewards (ibid., Article 4).

Performance measures and targets

The performance measures in the national police covenant include output measures (fines, charges), subjective performance measures (customer and citizen satisfaction with police work), and measures of internal performance (timeliness, efficiency, sickness absence, quality of telephone service). Based on these measures, targets will be set for each of the regional forces. Table 6.1 provides an overview of the measures and goals included in the covenant (BZK en Justitie, 2003).

The financial rewards

The annual budget earmarked for performance rewards is EUR 52 million (BZK, 2002d). This is equal to 2 percent of the total police budget; 98 percent of the budget will still be allocated based on exogenous factors such as miles of roads in the region. Of the total performance budget, 75 percent is used for rewarding superior individual performance of a force (the 'individual component') and 25 percent for rewarding superior relative performance (the 'benchmarking component'). Each force will receive a reward equal to 1.5 percent of its budget if it meets the performance targets for the period 2003-2006. The reward will most likely be allocated on an annual basis. Additionally, there is a (relatively small) reward for performing better than other

²⁵ BZK (2002d): "Despite all efforts of the forces and the audit that I conducted, I have to acknowledge the fact that we have not succeeded in achieving a level of data quality that allows legitimate payments. The latter point, the legitimacy, makes paying out the relative performance reward impossible. Therefore, I have decided to pass payments for another year." Translated from: "Ondanks alle inspanningen van de korpsen en de door mij uitgevoerde audit, heb ik moeten constateren, dat het nog niet goed gelukt is de kwaliteit van de gegevens op een zodanig peil te brengen dat een rechtmatige betaling mogelijk is. Dit laatste punt, de rechtmatigheid, maakt de uitbetaling van de benchmark component onmogelijk. Derhalve heb ik besloten ook dit jaar niet tot uitbetaling over te gaan".

forces of 0.5 percent of the police budget. These relative performance rewards make up only 25 percent of the total budgets for rewards. At the time of our writing, it is unclear how the benchmarking contest will be designed.

Table 6.1 Performance measures and goals for 2006 in the national police covenant

Performance measure	Goal
Output measures	
Number of fines and transactions	180,000 additional fines and transactions after police stops per year - an increase of 15 percent
Number of charges forwarded to the public prosecutor	40,000 additional charges forwarded to the prosecution council annually - an increase of 20 percent
Subjective performance measures	
Satisfaction with police services	Substantial improvement in percentage of citizens that is (very) satisfied with the most recent contact with the police as measured in the Citizen Police Monitor (PMB), a substantial improvement is achieving the highest score attained by that force over the period 1993-2002
Satisfaction with availability of police	Substantial improvement in availability as measured in the PMB survey, a substantial improvement is achieving the highest score attained by that force over the period 1993-2002
Measures of internal performance	
Timely processing of charges against young, persistent offenders	80 percent of the charges against (young) persistent offenders should be forwarded to the public prosecutor within 30 days
Overall efficiency gains	Efficiency improvement of at least 5 percent in terms of active police officers - forces that have a sick leave percentage of lower than 8 percent can subtract the difference from the 5 percent efficiency improvement
Sickness absence percentage	Sickness absence percentage should be reduced from 10 to 8 percent
Quality of emergency and other telephone service	To be defined goals for responsiveness to 0900-8844 calls (National Phone number Police, Landelijk Telefoonnummer Politie), responsiveness to emergency 112 calls, and the overall quality of 0900-8844 calls

6.4 Conclusions

The performance contracts are meant to improve the forces' accountability and incentives in light of the agreements included in the new 'National Police Covenant'. Each force will receive a modest financial reward if it meets the performance targets. The contract includes output measures (fines, charges), subjective performance measures (customer and citizen satisfaction), and measures of internal performance (timeliness, efficiency, sickness absence, quality of telephone services). Other measures may be included. Additionally, there is a (relatively small)

reward for performing better than other forces. At the time of our writing, it is unclear how the benchmarking contest will be designed.

7 Anticipating gaming of the performance evaluation

7.1 Aim and structure

The Department's challenge is to design performance contracts in a way that results in well-targeted incentives. Success is not guaranteed. Recall the frustration of London police officers caused by apparently inappropriate performance measures (see Chapter 5).

Performance contracts may not only fail to send the right signal to the police, but may also invite the forces to distort the signals sent back to the Department. This strategic response to performance evaluations is called 'gaming'. Gaming stands for all behaviour that allows the forces to boost their measured performance but is inefficient from the perspective of the Department's objective for the forces. Gaming includes manipulation of performance data, but also undesirable changes in the way the work is done. Think of a police officer who fines just about everything in the last weeks of December to meet the annual target number of fines.

Clearly, the Department will need to anticipate on a strategic response when choosing a method of performance evaluation. In this chapter, we identify the dilemmas related to gaming of performance evaluations. That will help us to better understand the difficulties facing the Department.

7.2 Objective performance measures and the tradeoff between incentives and gaming

Ideally, the Department would hold the forces accountable to outcomes. Outcome measures such as crime levels and feelings of safety reflect those things that we want to see from the police. By only including outcome measures in the performance contract, the Department leaves it to the forces to organize their internal processes and choose a mix of outputs in a way that contributes most to the desirable outcomes. Such a policy would be in line with the desirable discretion of the regional police forces (see Chapter 2).

For example, if crime prevention was all the Department cared about, then the ideal measure would tell us how many crimes have been prevented by the police in that region. This figure would be related to the force's budget to compare its performance with that of other regions. With such a measure, incentives are perfectly aligned with the objectives for the forces. Moreover, the measure would be under perfect control of the forces. A change in the measure would unambiguously signal a corresponding change in the performance of the force.

Why the ideal measure is not available

Unfortunately, outcome measures often do not provide a good picture of police performance. They tend to be affected by many factors and actors other than the police. Feelings of safety are a good example. A terrorist attack may be a more important determinant of feelings of safety than the any policing strategy, see De Mulder (2002).²⁶ Identifying the contribution of the police to a change in outcome measures is often unfeasible. For example, attempts to identify the effect of police work on crime for a country the size of the United States have not been very successful (see McCrary, 2002), making it virtually impossible to identify an effect at the level of a Dutch police region. The changing nature of policing also makes it more difficult to evaluate its outcomes (Travis and Waul, 2002, p. 19). The new commitment to engage in a wider variety of partners in the development of 'safety plans' makes it harder to attribute any successes to the work of the police.

Why 'noisy' outcome measures do not provide much incentive

Including such 'noisy' outcome measures in a performance contract would not result in strong incentives for two reasons. First, making the reward dependent on a measure that is not under close control increases the risks for forces. Second, such measures increase the information asymmetry between the forces and the Department.

(a) 'Noisy' measures imply higher risk

Noisy performance measurement makes performance contracts more look like a lottery than a reward for effort. For example, if we notice an improvement in crime levels over one year, the improvement may or may not be due to the police. If we cannot distinguish between noise and performance, then performance measurement is not likely to strengthen incentives. After all, when the return to efforts to fight crime are highly uncertain, the police may decide it is not worth putting in high effort (assuming risk aversion, see Prendergast, 1999, Section 2).

Offering a higher reward to offset the uncertainty does not fully compensate for the loss in incentives. At first sight, a higher reward may induce a force to take the risk and put in high effort. High rewards make a force's budget strongly dependent on external shocks, however. The Department may have to let the forces build up some buffer capital to absorb the shocks in the noisy measures.²⁷ But that goes at a cost of the power of incentives.

²⁶ See Van Dijk et al. (2000) for empirical research into the determinants of feelings of safety.

²⁷ Buffer capital will work as long as shocks are truly randomly distributed over time and across forces. In that case, the noise in the indicator tends to cancel out. Buffer capital will not suffice when some forces are dogged by bad luck or when all forces are struck by bad luck for a long period of time. To handle these cases, a specific force may receive additional funds or all forces may receive additional funds to overcome difficulties. Clearly, such a bail out goes against the idea of rewarding results.

(b) 'Noisy' measures make it easier to hide behind external factors affecting performance

When performance measurement is noisy, the information asymmetry between the Department and the forces tends to increase. The forces could use this asymmetry by referring to 'extraordinary circumstances' when performance is low (Blumstein, 1999). The forces may ask the Secretary to overrule the conditions specified in the contract with the argument that some external actor or factor negatively affected the performance measure. Clearly, the possibility of hiding behind unfavourable external conditions lowers the power of incentives.

Measures of output and internal performance as substitutes

Given the low incentives resulting from noisy outcome measures, the Department could look for measures of output and internal performance that are highly correlated with the desired outcomes. For example, the number of breathalyser tests for drivers could contribute to the outcome of greater traffic safety. The number of tests is a measure the police can do something about. It is a *controllable* performance measure. There is a clear link between effort and performance. In other words: the power of incentives is higher.

Clearly, such measures have to be used with caution since they intervene in the 'internal machinery' of the forces. By using output measures and measures of internal performance, the Department prescribes the best way of achieving the outcomes. Measures of internal performance intervene even deeper in the internal optimization processes than output measures. In both cases, the information asymmetry between the Department and the forces makes it hard to write such 'prescriptions'.

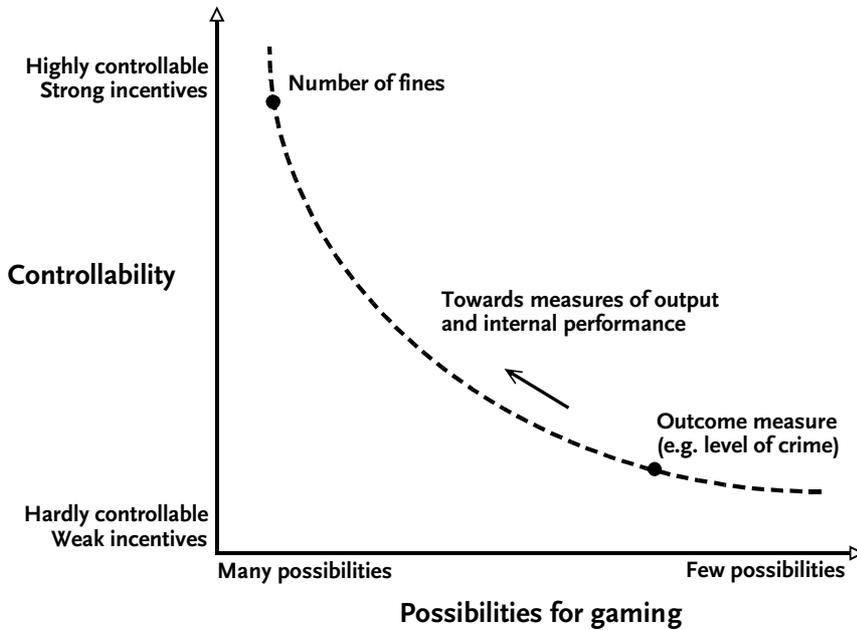
Moreover, the forces may take advantage of this information asymmetry by looking for the easiest way of meeting the Department's wishes (so called gaming). In police work, the easiest way of producing output or increasing internal performance is often not the best way. Unsurprisingly, experiences in New Zealand show that the easiest way of increasing the number of breathalyser tests is not likely to be in line with the objectives for the police. At the end of 1998, Wellington traffic police had carried out nowhere near the number of breathalyser tests that it was contracted to perform. To meet its contractual obligation, the police brought Wellington to a standstill, breathalysing almost everyone in a motor vehicle (UK Parliament, 1999). Another example of a performance measure that can easily be gamed is the number of fines (to be discussed in the next chapter).

Baker's principle

Baker (2002) addresses this regularity between controllability and gaming (we abstract from data manipulation for now, we deal with this issue later in this section). Based on this article, we formulate Baker's principle: *When holding an organization accountable to performance measures*

(i.e. when keeping the incentive structure constant), more controllable performance measure - that provide stronger incentives - are more likely to result in undesirable behavioural effects. Many interviewees emphasized the relevance of this principle to the case of policing. Figure 7.1 illustrates the tradeoff graphically.

Figure 7.1 Baker's principle: the tradeoff between controllability and gaming



How could controllable measures reward the wrong behaviour?

In the case of policing, controllable measures of output and internal performance only reflect some dimensions of police work. When only some of the efforts contributing to the outcomes of police work are rewarded, the forces gain from substituting away from tasks that do not produce measurable results to tasks that do.²⁸ After all, if the incentives are effective, then the forces are more likely to work on tasks that do contribute to the performance measures than on tasks that do not.

Low measurability of the output of some activities is one of the reasons why performance measures do not capture all of the dimensions of police work. Such measures can lead to a bias towards activities that can be quantified at the expense of activities that cannot so easily be quantified (Holmstrom and Milgrom, 1991). And even when activities produce measurable

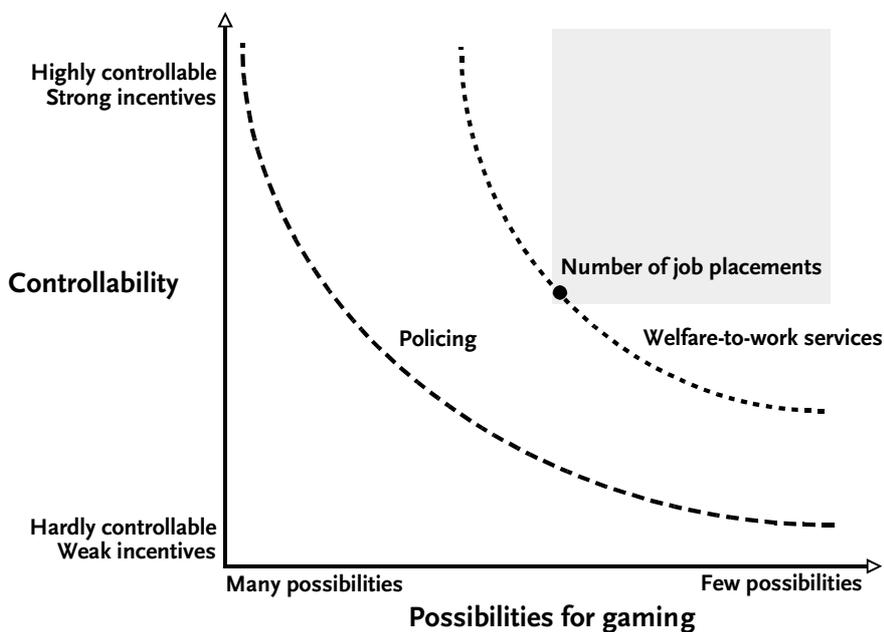
²⁸ Assuming that substitution between tasks is possible. If tasks are complementary, an incomplete contract will not create a bias in police work.

outputs, the performance measure may not cover all dimensions of output. Consequently, the forces could be tempted to shirk on quality of their services (Hart, Shleifer and Vishny, 1997).

How does the police compare to other sectors?

Capturing all of the essential dimensions in objective performance measures is particularly challenging in the case of policing. Not only is the police engaged in many activities ranging from assistance at traffic accidents to resolving a hostage stand off, each individual activity involves many different dimensions as well. The problem is more complex than in the case of, say, welfare-to-work services (see Vollaard, 1999). These services are aimed at an easily observable outcome: placement into jobs. The tradeoff is still valid. Performance measures that account for the quality of the job placement tend to be less controllable. For example, the duration of a job is not completely under control of a job consultant. However, the tradeoff between the power of incentives and possibilities for gaming will be more favourable in the case of welfare-to-work services than in the case of policing. Figure 7.2 illustrates the difference between the two sectors. Say that the shaded part shows the allowable area from the government's perspective. Within this area, an agency has reasonably strong incentives without all too many possibilities for gaming. Some performance measures for welfare-to-work services are within the shaded area, whereas there are no such measures for policing.

Figure 7.2 Comparing policing to welfare-to-work services



Data manipulation as additional problem

Measures the police ‘can do something about’ tend to be more often administered by the police themselves. Controllable performance measures are even less attractive if they are prone to manipulation of data. With strong incentives, it is very tempting to present a rosier picture of the situation. Clearly, close control over data collection within the force is a precondition for successful manipulation. Relative performance measurement creates a particularly strong incentive to manipulate data. After all, it is hard to stay behind if everyone else is manipulating its data. Why miss out on a reward just because another force is showing unethical behaviour? As we will discuss in Chapter 8, there are ways to (partly) get around the problem of data manipulation (such as relying on third parties for data). Therefore, data manipulation may not be as much a driver of the ‘inescapable’ tradeoff between incentives and gaming as the multi-dimensional nature of police work.

Police data, crime statistics in particular, are known to be highly manipulable. Methods to manipulate crime data include (a) influencing the inclination to register crime by showing a lack of responsiveness, (b) not recording a crime allegation (Bovenkerk, 2002, p. 11, 12)²⁹, and (c) creative counting.³⁰ The ‘crime mix’ can be manipulated by selectively recording crime allegations and altering the classification of reported crimes.

An example of the dilemma of data manipulation is the moral outrage of David Swift, Deputy Chief Constable of the British force Staffordshire. His force ended third worst in a league table published by ‘The Observer’. However, Swift claimed that Staffordshire’s low ranking was not due to low performance but to operating ‘the most ethical crime recording system in the country’ (Staffordshire Police, 2001).³¹

Manipulation is not limited to crime statistics, however. For example, the Amsterdam police tried to bring down their sickness absence rate from 13.0 percent - the highest of all police forces - to a more favourable 6.9 percent (*Het Parool*, January 6, 2003). When the Interior Department checked the reported absence rate, it found that the definition of a sick person had been altered.

²⁹ Not all crime allegations end up as crime report. After all, in many instances police officers will not be content to file a crime report without clear supporting evidence that the offence has been committed. Even if there is reasonable evidence that an offence has been committed, but the complainant refuses to press a charge (as in many cases of non-stranger violence), then there is a general reluctance to make any crime record at all (Burrows et al., 2000, p. 68).

³⁰ Sherman (1998) mentions the removal of three commanders at the New York City police over a five-year period for improperly counting crime to artificially improve their performance.

³¹ He continues: ‘We could dramatically reduce our crime rates overnight if we chose to operate recording systems similar to a large number of other forces who, on the face of things, appear to have a rosier crime picture than us.’

Contrary to the policy followed by other forces, long-term ill people and people who are partly disabled had been excluded.

7.3 Shifting from objective to subjective performance evaluation

Given the multi-dimensional nature of police work and the importance of quality of police work, the tradeoffs related to the available objective performance measures may not provide for satisfactory choices in the case of policing. If this is indeed the case, then the Department may have to look into subjective performance assessment. As we will discuss below, subjective performance assessment allows a better position on the tradeoff between incentives and gaming.

Subjective assessments can either be conducted by experts or by customers. Peers are the best experts. As described in Chapter 3, peers are involved in audit teams and review committees that assess quality management of forces. They could critically evaluate the approach taken by a force. Clearly, their assessment cannot be translated in some quantitative performance measures (more on this issue in Chapter 10). Customers include citizens and businesses, but also the prosecution council. For example, in New Zealand, Judges' satisfaction with police services is measured on a yearly basis (New Zealand Police, 2002).³² In the Netherlands, citizens' perceptions of police are surveyed bi-annually (Citizen Police Monitor, PMB).

Why would subjective performance assessment lead to better-targeted incentives?

The attraction of subjectively determined measures of performance is that they allow a more complete picture of performance to be attained, not possible with objective measures. By including more of the dimensions of the work into the evaluation, subjective assessment limits possibilities for gaming. Given its multi-dimensional nature, this feature is of particular relevance to evaluating police work.

Moreover, in light of the tradeoff discussed in the previous section, subjective performance measures incite the right response without losing all the power of incentives (as tends to happen in the case of outcome measures that are in line with the Department's objectives). For example, businesses state that the determinants of their satisfaction with police services include speed of recording a crime, the capability to solve the problem at hand, and courtesy in contacts with the police (Visser et al., 2002b, p. 64). These are all things the police can do something about. There may not always be one-to-one link between efforts and performance. Customers may have difficulty assessing a monopolist. They cannot easily compare the performance of several police

³² The 'judicial satisfaction survey' include indicators of fairness, appropriateness of charges, sufficiency of evidence, presentation of evidence and knowledge of the law and advocacy.

forces, making it difficult to assess their performance. Very poor performance or very high performance is likely to be picked up in the evaluation though. Assessment by citizens may also be somewhat noisy if their perceptions are influenced by media reports about police actions elsewhere.

A commonly observed response to subjective assessment, gaming of the evaluation by either the evaluator or the evaluated party, may not be as problematic in the case of policing:

- *Gaming by the forces.* Inefficient behaviour just to please the evaluator does not seem to be much of a problem (think of the salesman who gives too much price brakes ‘to please customers’, so called ‘rent-seeking behaviour’, see Prendergast, 1999, section 2.2.3). For example, efforts aimed at the abovementioned drivers of satisfaction stated by businesses would be in line with the Department’s objective of ‘efficient, effective, fair and decent policing’. Moreover, subjective measures are not susceptible to data manipulation. By definition, the Department relies on third parties for performance data.
- *Gaming by the evaluators.* A number of problems related to gaming by evaluators have been highlighted in the literature, ranging from ‘leniency bias’, where evaluators are reluctant to give bad ratings, and ‘centrality bias’, where evaluators compress ratings around some norm rather than truly distinguishing good from bad performance (ibid.). The assessments by citizens or public prosecutors are not likely to be tainted by such hidden agendas. They are likely to give honest responses. It may be more of a challenge to guarantee objectiveness of experts when assessing the police (think of the peer review briefly discussed in Section 3.2). After all, experts tend to be not only well-informed but also well-known in the field. If the peer review is not properly designed, social pressure could possibly result in biased evaluations.

Another common problem with subjective assessment - inherent biases in the perceptions of evaluators - does not seem to be prohibitive either. Evaluators may not take all dimensions of police work into account. For example, citizens may be biased towards visible aspects of policing such as street patrols. Therefore, using citizen surveys to assess overall police performance may lead to an undervaluation of not-so-visible police work. An indication of such a bias is the claim of some criminologists that the current emphasis on ‘visible blue’ (‘blauw op straat’) is not justified given the under-investment in criminal investigations (see Boom, 2002). By limiting the assessment to only those parts that are relevant to that specific group of customers, such biases can be limited. Additionally, since citizens are not directly faced with the costs of police work, they may not be fair evaluators. Someone who directly pays for a service will weigh costs and benefits. However, someone who does not face the marginal costs may put an unrealistic demand on police services (see also Chapter 2 on the need to ration the supply of

police services). Part of this bias may be overcome by comparing citizen satisfaction ratings over time and/or across regions. Benchmarking provides the way out.

7.4 Conclusions

Holding the police forces accountable to results may not be as easy as it sounds. Outcome measures such as the level of crime are not quite measures the police can do something about. Using measures of output and internal performance that are correlated with the outcomes instead may lead to undesirable effects. The forces may inefficiently focus on 'meeting the numbers' at the cost of other police work and the quality of output. In the case of policing, there do not seem to be objective performance measures that both invite the right behavioural response and are also controllable. Therefore, the Department may wish to rely on more subjective assessment of police performance by customers and peers. Subjective assessment includes many more dimensions of police work in the evaluation. The police can also do something about the appreciation of their work by customers and peers. Keeping these dilemmas in mind, we will have a closer look at the proposed performance contracts in the next chapter.

8 Effectiveness of the proposed performance contracts

8.1 Aim and structure

Now that we have described the outlines of the performance contracts and the difficulties related to evaluating police performance, we will assess whether the contracts are likely to be effective. We judge the effectiveness by the objectives stated in Chapter 2: the forces should reduce crime, promote safety and reduce disorder in an efficient, effective, fair and decent manner. In the following three sections, we will look into three dimensions of contract design: (1) performance evaluation, (2) collection of data, and (3) design of the performance rewards. Instead of ending each individual section with a set of conclusions, we will bring together all the findings in the concluding Section 8.5.

8.2 Assessing the proposed method of performance evaluation

How does the Department deal with the dilemmas in evaluating police performance discussed in Chapter 7? In the first two sections, we discuss the likely behavioural response to the evaluation of efforts to reduce crime and to promote safety and reduce disorderly behaviour. Then we deal with measures aimed at both objectives. Before assessing the individual measures, we note that there is no indication that the total number of measures is excessive in the sense of creating a lack of focus for management.³³

8.2.1 Measuring efforts to reduce crime

There are two performance measures related to the aim of reducing crime included in the performance contracts (see Table 6.1):

- One output measure: *number of charges forwarded to the public prosecutor per year.*
- One measure of internal performance: *timeliness of forwarded charges against (young) persistent offenders.*

³³ There is no hard evidence on the number of measures management can handle. The measures used for own purposes by a district chief of police was not smaller (interview Marc Jacobs, see Annex 2) - although this practice may be partly the result of the Department's requirements.

Reducing crime is made synonymous to the number of charges produced by a force.³⁴ Both measures focus on efficiency. The more efficient a force operates, the higher the number of forwarded charges and the faster charges are being forwarded to the prosecution council. Implicit assumption is that being efficient in these two tasks implies higher effectiveness of regional police work.

The chosen output measure, the number of charges, is under close control of the forces.³⁵ To increase the number of charges, the police can allocate more resources to criminal investigations or try to work more efficiently. In line with our argument in Chapter 7, strong incentives to increase the number of charges could compromise the quality of police work (possibilities for data manipulation are limited by relying on the prosecution council for data).

Less importance attached to sound evidence?

The forces may be induced to also forward cases with rather poor evidence to the public prosecutor (see 'Afrekenen op prestatie leidt tot bureaucratie', *NRC Handelsblad*, January 3, 2003). The performance contract does not include a measure reflecting the quality of evidence to redress such a bias.³⁶ A subjective assessment of police performance by public prosecutors is not part of the performance contract either. There are countervailing powers that put a restraint on gaming, however. The larger offices of the prosecution council second employees to police stations to check the quality of the charges ('Secretaris op locatie'). Cases without insufficient evidence are filtered out. The temptation to game the evaluation remains, however. The (financial) incentives are based on the measures of the Interior Department after all.

Possible bias towards easiest charges

A focus on the easiest charges inflates performance, but is not an efficient way to reduce crime. This bias could be limited by attaching weights to charges that reflect the severity of the case.

³⁴ Ironically, the more crime, the more charges, and the better the police seems to be able to perform. The number of charges is less influenced by external conditions than one might think, however. Because of a lack of resources, there is an excess supply of cases that could probably be solved but are 'put on the shelf' (the 80,000 'plankzaken', see TK27834, nr.2). If the police is able to shift resources to criminal investigations, then the number of charges forwarded to the public prosecutor is under control of the police - at least within certain bounds. Actually, the target of 40,000 additional charges in the national police covenant is directly based on an estimate of the number of 'shelved cases' that can be solved.

³⁵ The less controllable outcome measure 'feelings of safety' that was part of the system of performance rewards (see Annex 1) has not been included in the performance contracts. During the contract negotiations, the police forces indicated that they did not want to be held accountable to a measure that they do not have under control (interview Department of the Interior).

³⁶ In the system of performance rewards, there was the measure 'percentage of charges dismissed because of inconclusive evidence' (see Annex 1). This measure has not been adopted in the national police covenant.

The output measure 'weighted solved crimes per officer' introduced by Wiebrens (2002) does just that (see Section 4.4). No such weights are used, however. Maybe the forces did not want to be held accountable to a measure that is somewhat more noisy. After all, the prosecution council also has influence on the conviction of charged individuals (the council needs to ensure sufficient capacity in the courts and the public prosecutors need to do their work well too). There are other countervailing powers that may limit gaming, however. The prosecution council requires the forces to use a tool that indicates the severity of the offence and the likelihood of solving the case (Openbaar Ministerie, 2003). With that tool, the council decreases the information asymmetry - and therefore increases its say over the crime fighting efforts of the police.

Fair and decent policing out of the picture

The number of charges does not reflect the objective of fair and decent crime fighting (neither does the measure for timeliness). Apparently, the Department does not think that the incentive to increase the number of charges may go at a cost of fair and decent policing. This policy is at odds with practices in Australia and the United Kingdom.³⁷ It is uncertain whether this is a deficiency within the Dutch context.

Undervaluation of pro-active policing

The police does not only produce charges against offenders, it is also the police's objective to prevent crimes from occurring. Indeed, the risk of getting caught prevents crime, but it is not the only way of preventing crime. There are also so-called 'pro-active approaches', including patrolling the streets and crime prevention campaigns (watch out for pickpockets, lock your doors, etc.). These activities are neither reflected in a greater number of charges nor in other performance measures included in the contract. Consequently, pro-active policing may be undervalued.

Prescribing improvements: timeliness of charges against (young) persistent offenders

In addition to one output measure, the contract also includes a measure of internal performance: the timeliness of forwarded charges against (young) persistent offenders. The Department is convinced that faster processing of these specific charges is the most expedient way of increasing the effectiveness of police work in all of the regions. More specifically, the Department assumes that (a) young, persistent offender should be given priority over other cases and (b) timely processing of charges is specifically important to reduce crime. We do not know the specifics of the empirical evidence that could justify this policy (better crime

³⁷ For example, Australia includes 'death in police custody and custody related operations', 'complaints per 100,000 people and per 100 sworn police staff' to complete the picture of police performance (see Chapter 9).

prevention does not seem to be a good argument for this policy, see Bosker, 1997). Apparently, the other performance measures are thought to provide too little incentive to focus on (young) persistent offenders. We did not find indications that the forces can easily game this performance measure. If speeding up charges is the best way of improving police performance, then this performance measures may indeed contribute to better policing.

8.2.2 Measuring efforts to promote safety and reduce disorderly behaviour

There are several performance measures related to the aim of promoting safety and reducing disorderly behaviour (see Table 6.1):

- One output measure: *number of fines per year*.
- Several related measures of internal performance: *quality of telephone services*.
- Two subjective performance measures: *citizen satisfaction with police contact and police availability*.

In this section, we will discuss each of the types of performance measures.

Frivolous fining

The number of fines is the ultimately controllable performance measure.³⁸ As we already indicated in Chapter 7, this measure also invites gaming. Indeed, the forces may inefficiently increase the number of fines in two ways. First, the police may be induced to substitute warnings and other informal ways of resolving situations with fines. Such a substitution may not necessarily be desirable. If it were desirable, then such a decision should preferably be made at the local level. After all, police work relies heavily on the discretionary power of the officers to resolve situations (see Chapter 2). Second, police officers could fine ‘the easy way’ to inflate their performance. For example, an additional check on seat belt use may be a low-cost solution to inflate performance. The cheapest way of fining is not likely to be the most desirable way. Therefore, the number of fines is not likely to contribute to well-targeted incentives aimed at promoting safety and reducing disorderly behaviour.

Telephone services

The Department also includes measures of internal performance reflecting the quality of emergency and other telephone services. The Department is convinced that improving quality of telephone contact is an expedient way of promoting safety and reducing disorderly behaviour.

³⁸ The noisy measure ‘quality of living in the neighbourhood’ which was part of the performance rewards (see Annex 1) has not been included in the national police covenant (for empirical research into the determinants of the perceived quality of living see Van Dijk et al., 2000).

We do not know all of the empirical evidence, although evaluations have shown serious shortcomings in the service in the past, see Nationale Ombudsman (2001). The measures does not seem to allow for much gaming. Therefore, we conclude that this measure could be a useful way of creating a more complete picture of police performance.

Rewarding citizens' satisfaction with police services

The reliance on subjective assessment by citizens is not surprising given the relatively low measurability of the outputs of community policing and emergency services. As we discussed in Chapter 7, subjective performance measures provide relatively well-targeted incentives without losing all the power of incentives as happens in the case of outcome measures.

8.2.3 Measuring efforts directed at both objectives

There are two measures of internal performance that contribute to both objectives: the sickness absence percentage and overall efficiency savings.

Sickness absence

The rate of sickness absence reveals something about the efficiency and the 'health' of an organization. The focus on sickness absence is supported by the current level of sickness absence of about 10 percent and the fact that manpower accounts for some 75 percent of total costs of the police (some British police forces have a sickness absence percentage of about 6 percent, for an overview see Arnott and Emmerson, 2001). Additionally, Verhoeven et al. (2002, p. 135) find that the police invests little in lowering the rate of sick leave compared to the private sector. Given the current levels of sickness absence, these low investments are not justified by great returns to effort. Therefore, the focus on sickness absence may well be justified. Clearly, the importance of this issue also depends on the starting point of a force: if sickness absence is already low, then further decreases may be very costly. The Department seems to built in a floor of 8 percent. If a force goes below the 8 percent, then it can subtract the difference from the required efficiency improvement of 5 percent (see Table 6.1).

Overall efficiency savings

The measure 'overall efficiency savings' is a direct indicator of improvement in efficiency. Targets for efficiency savings are force-specific: the required efficiency improvement is dependent on their success in reducing sickness absence. An efficiency savings operations explicitly mentioned in the national covenant is better resource management (mainly scheduling manpower). This is a logical suggestion given the recent findings of the General Accounting Office (Algemene Rekenkamer, 2003a, p. 25, 26). They find that current resource management of the forces is not based on any analysis of crime and other problems within their region. It is unclear how these 'overall efficiency savings' will be measured, however. Activity measurement

is underdeveloped. It will be hard to show how, for example, better resource management has resulted in increased efficiency. Without better activity measurement the efficiency savings may be based on guess work, and therefore be susceptible to manipulation.³⁹

In this section, we assumed that the data are reliable (apart from manipulation) and that targets and rewards are set right. In the following two sections, we will explore collection of data and the setting of targets and rewards.

8.3 Assessing collection of data

Without proper data, performance contracts cannot be implemented. Clearest evidence of the importance of good data is the failure of the performance rewards discussed in Chapter 6. The system of performance rewards never got off the ground due to measurement error (although the system may have been stillborn for other reasons as well). In this section, we discuss the Department's efforts to collect the data supporting the new performance contracts. We discuss how measurement error weakens incentives first.

Measurement error weakens incentives

Measurement error weakens incentives just like noisy performance measurement does. Non-uniform and incorrect collection of data also introduce noise in the performance measures. In that case, the forces are held accountable to measures that do not fully reflect their performance. Consequently, the risk that high effort does not result in higher performance (and a reward) blunts incentives (assuming risk aversion). Moreover, measurement error allow the forces to hide behind 'technical and administrative issues' when their performance is sub par. In its extreme form, measurement error will make any performance evaluation impossible (as has been the case over the last years).

Measurement error may also lead to frustration, which undermines the intrinsic motivation to improve performance. For example, police force Rotterdam-Rijnmond reacted angrily when differences in counting crimes negatively affected its place in the crime clear-up rate ranking published by Netherlands Statistics (CBS).

Problems concentrate in police data

Given the patchwork of information systems across police regions and the sorry state of these information systems (KBB, 2002), measurement error is particularly problematic for police

³⁹ The British authorities were facing a similar problem. As a solution, they summoned the Inspection to verify claimed efficiency savings through detailed assessments (Home Office, 2002).

data. On a scale of one to ten, the forces mark reliability of their data six (In 't Veld et al., 2001, p. 87). Cited causes of data unreliability include lack of user friendliness of the recording systems and lack of uniformity to properly enter the facts. To indicate the size of the problem: the forces stated that recorded crime would be reduced by 10 percent when police data would be properly cleaned (Algemene Rekenkamer, 2003a, p. 33).

Unreliability of police data has been an issue for a long time. Algemene Rekenkamer (2003a, p. 36) notes that the very same problems of data quality were identified in a similar study in 1995. The structural unreliability of data is another indication that performance improvement has little priority. After all, best practices can only be identified when keeping track of operations. Actually, the system of performance rewards created perverse incentives for the forces not to improve their information systems. Low data quality did not allow for legitimate payment of the performance rewards - but the forces received the performance rewards anyway.⁴⁰ The national police covenant has only mention of future agreements between the Department and the forces on the data quality issue without mentioning deadlines - although expectations of firm agreements were created in BZK (2002d).

Concerns limited through reliance on third parties

To circumvent the problems with police data, the Department relies as much as possible on third parties for performance data, including the prosecution council and the Centraal Justitiele Incassobureau (Central Collection Agency). The subjective assessment of police performance is based on a population survey, which is conducted uniformly across the country. There still are a few items that are based on police data, including the sickness absence percentage and the quality of telephone services. Apart from some concerns about manipulation (think of the artificially low rate of sickness absence in Amsterdam discussed in Section 7.3), we have no indication that non-uniform and incorrect collection of data pose major problems for these items. Therefore, measurement error does not seem to undermine the incentives resulting from the performance contracts.

8.4 Assessing the proposed design of the performance rewards

Based on the performance measures, the Department defines targets that the forces should achieve to get the performance reward. There are three common problems related to setting targets and rewards: (1) individual targets are set too high or too low, (2) comparing and

⁴⁰ BZK (2002d, p. 6): "The non-allocated performance budget of 2002 (EUR 10m.) has already been spent on increases in manpower which are part of the 2003 budget." Translated from: "Het vrijvallende prestatiebudget in 2002 (EUR 10 mln.) is reeds opgenomen in het budget 2003 voor sterkte-uitbreiding".

rewarding performance *relative* to other forces (so called yardstick competition), and (3) the chance that strong financial incentives put low performers in a negative spiral.

8.4.1 **Setting performance targets**

Performance targets are threshold values of performance measures above which a force is awarded the financial bonus. There are some serious challenges to setting targets and rewarding efforts in the case of policing. After discussing these challenges, we assess the proposed targets.

First challenge: setting targets at the right level

Setting targets at the right level requires knowledge of the way the police operates and knowledge about the environment in which they operate:

- *Information about internal processes.* For example, the Department has to judge whether it is reasonable to demand 40,000 additional charges to be forwarded to the public prosecutor. Such information will be hard to obtain, since the forces may not have the data themselves. Activity measurement is underdeveloped and so is resource management (Algemene Rekenkamer, 2003a). Given the general lack of good information on business processes within the police, it will be very hard to set the targets right.
- *Information about the environment.* Targets need to be related to reality. The contracts deal with problems like crime that are to a great extent driven by forces external to the Department and the police. Without a clear connection to reality, police work would be supply-driven.

Whether targets strengthen incentives depends on whether they are set just right, too low or too high. If the target is set right, it is within reach of the force that puts in high effort. In that case, high effort will lead to an appropriate reward. If the target is too low, the force may stop putting in high effort as soon as the target has been achieved. If the target is too high, the force may decide not to put in high effort since the reward is not within reach anyway. Given the lack of good information discussed above, it is unlikely that the target will be set just right. Whether a target is set right also depends on the size of the performance reward. After all, if the cost of additional effort exceeds the reward, the force has no incentive to put in high effort.

Second challenge: handling multiple targets

Clearly, the Department does not set one single target. With multiple targets, the problem is how to weigh each measure. Is there a specific reward for meeting every single target? If so, is the reward equal for all the targets? Or is there a rule how many of the targets have to be met as condition for receiving the reward? The system of performance rewards had a rather arbitrary way of allocating the performance rewards. The forces had the flexibility to do well on the

measure of choice. Moreover, no weighting was applied to any of the measures. Clearly, biases in police work may result if the forces will choose to only work hard on the easiest targets.

The targets in the national police covenant

At this time, we only have information on the overall targets for the individual measures for the period 2003-2006. These national targets will have to be transformed to region-specific targets and, maybe, to year-on-year changes. Therefore, we will focus on the national targets for now. One target that seems to be set rather low is the number of fines and transactions. The target is to write 180,000 additional fines annually by 2006. This target implies some 19 additional fines annually per active police officer, which is not a major achievement.⁴¹ The 40,000 additional charges do not seem to be much of a challenge either - considering the increase in police officers over the years to come.⁴² When these targets have been set too low, then they will not provide much incentive. The goals for customer satisfaction may also be little ambitious since they are based on historical performance ('perform at least as good as you did in the past'). We have too little information to assess the other targets in the police covenant (such as timely processing of charges and quality of telephone services). At this moment, it is unclear how the Department will weigh the different performance measures to allocate the individual performance rewards.

With a fixed number of 40,000 additional charges and 180,000 additional fines per year, an implicit assumption is that the level of crime and disorderly behaviour will not change over the period 2003-2006. Clearly, these targets make regional police work more supply-driven than demand-driven.

8.4.2 Comparing and rewarding relative performance

The performance contracts include a 'tournament' between the forces for extra budget. There are two design problems related to comparing and rewarding relative performance:

- *Accounting for different operating environments.* When comparing the performance of forces, the Department should somehow correct for the different (changes in) external conditions that the

⁴¹ The police employs some 50,706 FTEs, 75 percent of which are active personnel. Some 50 percent of their time is spent on back office work, training, etc. About half of the remaining time is spent on community policing and emergency assistance (Algemene Rekenkamer, 2003a, p. 47). That leaves us with 9,507 officers who could fine people. Therefore, 180,000 additional fines implies some 19 additional fines per active police officer. To compare: Terpstra (2002, p. 50) mentions 160 fines per officer annually as a goal set by a local police department. In that case, 19 extra fines implies an increase of about 12 percent, which is in line with the 15 percent mentioned in BZK and Justitie (2003).

⁴² Back-of-the-envelope calculations learn that 40,000 additional charges would probably take about 500 FTEs (including management and support), whereas the number of FTEs will be increased by 4,000.

regional police forces face. After all, Groningen and Amsterdam-Amstelland will probably face different 'external shocks'. When such corrections are not made, the link between effort and performance is distorted and the power of incentives will decrease.

- *Making the financial incentives worthwhile.* It will be hard to set rewards at such a level that the tournament is actually interesting for the forces from a financial perspective. With many forces competing for a reward, the risk of not receiving the reward increases. The *expected* reward of putting in more effort than another force may therefore be rather low (assuming risk aversion). To compensate for that risk, the reward could be increased. However, high rewards could lead to high fluctuations in police budgets. These fluctuations are not desirable from the standpoint of the citizens.

At this time, it is unclear how the relative performance rewards will be allocated. The Department is still negotiating with the forces about the design of the tournament.⁴³ The only thing that is clear is that only 25 percent of the performance budget will be used for the tournament. Therefore, the relative performance reward may not constitute a major incentive from a financial perspective.

8.4.3 Strong financial incentives and the possibility of a negative spiral

Strong financial incentives may put the quality and continuation of police services in jeopardy. High-performing forces earn the rewards, low-performing forces do not and may end up in a negative spiral of lower performance, less resources, and so on. The negative spiral may be reinforced by a loss of motivation among the low performers. If they have little to no chance of winning the reward, then why would they put in high effort? In that case, performance contracts may lead to a class of failing police forces that have little motivation to improve. This is undesirable since citizens have no choice other than to consume the police services provided in their locality. This problem is known as the 'Matthew effect'.⁴⁴ In principle, even with small financial rewards, these incentives go against the lack of consumer choice. After all, if a force misses out on a reward because of poor management, citizens will pay twice. They cannot switch to another police forces other than by moving somewhere else after all.

⁴³ In the system of performance rewards that will be replaced by the performance contracts, each force was put into a cluster of five comparable forces (based on degree of urbanization). Whether this solution actually works is uncertain. It is questionable whether external conditions are strongly correlated with the degree of urbanization. For example, the forces Limburg-Noord and Drenthe are in the same cluster, whereas their regional conditions, such as changes in the rate of unemployment, may well differ (see Chapter 9 for the interesting way in which the United Kingdom has tackled this problem).

⁴⁴ After Matthew 13:12 and 25:29: "Unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath."

The allocation of only a small part of the police budget (2 percent) is dependent upon performance. The 2 percent of the budget seems to be the part of the budget that is not already fixed (interview Frits Vlek). Although these rewards are relatively small, if they tend to be awarded to the previous winners, a gap between high and low performers may start to appear. Therefore, the national police covenant emphasizes individual performance rewards rather than relative performance rewards. That way, the low performers also have a good chance to receive a performance reward. Therefore, it seems unlikely that the performance contracts will throw low performing forces in a negative spiral.

8.5 Conclusions

Bringing together the outcomes of the analysis, we conclude the following: *considering the way in which performance is rewarded, the proposed performance contracts are not likely to be uniformly beneficial to the realization of the Department's objectives for the police forces.* Why is it that this intention may not be fully realized?

- *The performance measures only touch on some of the dimensions of police work, inviting the forces to narrowly focus on 'making the numbers'.* Such a strategic response may compromise the quality of police work, i.e. those dimensions that are not reflected in the measures. The problems concern the output measures rather than subjective performance measures (citizen satisfaction) and broadly defined measures of internal performance (such as the rate of sickness absence). Crime fighting is made synonymous with producing criminal charges at a high rate. This incentive is an invitation to also forward cases with questionable evidence to the public prosecutor and to focus on the offences that are easiest to solve. Moreover, crime prevention may be undervalued since such efforts are not fully reflected in the number of charges or other measures. Similarly, increasing the number of fines is an invitation to 'frivolous fining'.
- *Strengthening incentives requires more of target setting than may be possible.* Given the lack of (reliable) information about the internal processes of the forces, setting specific targets for each individual force is simply too challenging an undertaking. For example, back-of-the-envelope calculations indicate that the targets for the number fines and charges seem to be too little ambitious - certainly given the further increase in the number of police officers. Clearly, low targets do not provide much incentive to the forces.
- *Additionally, setting output targets threatens to turn the police into a supply-oriented rather than a demand-oriented organization.* Clearly, fixing output levels at the national level goes against the forces' need to react to (changes in) their specific regional environment.
- *Instituting financial rewards for meeting targets places financial gain above the public's wishes, which likely results in biases in police work.* There is little chance that the financial rewards reflect the social value of each measured dimension of police work. With a smaller number of rewards than

performance measures, the measures need to be somehow weighted. There is simply no information to weigh the quality of telephone services against timely processing of criminal charges, for example. So far the Department has not been able to find a satisfying solution to this problem.

Given these shortcomings in the design of the performance contracts, the question is whether these problems can be fixed. Can performance contracts be made effective - if only we find the right performance measures, set the right targets and choose the right way of allocating the rewards? We will address this question in the third step of the analysis.

Step III. What is the value of performance contracts for the police?

In Chapter 9, we will provide a brief overview of the policies that are being pursued in Australia and the United Kingdom. This look across the borders is meant to provide ideas for alternative ways of achieving the Department's objectives. Maybe there are alternative approaches that do not have all of the drawbacks of the proposed performance contracts identified in Chapter 8. Based on the analysis in the foregoing chapters and the lessons from Chapter 9, we draw conclusions about the value of performance contracts in Chapter 10.

9 Brief overview of policies in the United Kingdom and Australia

9.1 Aim and structure

Increasing accountability and incentives without introducing undesirable side effects; this is the challenge facing policy makers around the world. The police may be one of the most challenging sectors to take on. There should be a lot we could learn from countries that have taken the lead in this area, including Australia and the United Kingdom. In this chapter, we will provide a brief overview of the policies followed in these two countries. Given the absence of empirical studies into the actual effects of the policies in these and other countries, the lessons concentrate around ideas for the design of policies. We will not be able to draw hard conclusions on what works and what doesn't. In line with the scope of the study, we focus on the interaction between the national (federal) level and the regional level.

9.2 Policies in England and Wales

Subsequently, we describe the organization of the efforts to measure police performance, the actual performance measures and the way in which the measures are being used.

Organization of the efforts

The National Policing Plan 2003-2006 provides the framework for policies towards the police. In this document, the Home Office, the British equivalent of the Interior Department, defines the overarching aims and objectives for the police service. Within these aims and objectives, the Home Office then defines 'Ministerial priorities' that reflect the 'Home Office Public Service Agreement'. The Public Service Agreement states what the public can expect from their money and who is responsible for delivery of targets. The Agreement for the police service requires a reduction in crime and the fear of crime and an improvement in the performance of all forces. The national targets are supported by Best Value Performance Indicators (BVPIs).

Under the Local Government Act 1999, police authorities are responsible for securing Best Value in their region. Authorities must publish an Annual Policing Plan. The Plans provide a picture of the policing service delivered over the last 12 months; the authority's plans for next year; how the authority's performance has varied over time and compares with others; and how the views of local people have been taken into account. The Police Inspection (HMIC) reports publicly on the achievement of targets and compliance with acknowledged good practice.

Table 9.1 Best Value Performance Indicators**Corporate health**

Achievement of (2 percent) efficiency target as assessed by the Police Inspection (HMIC)
Annual estate running costs (excluding repairs and maintenance) per square meter
Number of complaints per 1,000 officers / Percentage of complaints substantiated / resolved informally
Percentage of the public satisfied with police action in response to 999 calls
Percentage of the public satisfied with the service received at police station inquiry counters
Percentage of victims satisfied with police initial response to a report of violent crime / burglary of a dwelling
Percentage of victims of road traffic collisions satisfied with the police service at the scene
Percentage of police officer strength which is female / of ethnic origin
Percentage of ethnic population of working age
Number of working days lost through sickness per police officer / civilian employee
The percentage of staff turnover of police officers / civilian employees
The percentage of police officers in operational posts / organizational support posts
Number of medical retirements of police officers (civilian employees) as a percentage of all officer (civilian) retirements
Actual net revenue expenditure per 1,000 population
Percentage of Police Authority buildings open to the public which are suitable for and accessible to disabled people

Crime

Level of crime / Fear of crime / Feelings of public safety
Public confidence in the criminal justice system or its component parts

Service delivery outcome

Total recorded crime per 1,000 population / The percentage detected recorded crimes
Domestic burglaries/Violent crimes/ Recorded robberies/Vehicle crimes per 1,000 households
Percentage detected domestic burglaries/violent crimes/robberies/vehicle crimes
Number of offenders dealt with for supply offences in response of class A drugs per 10,000 population
Number of public disorder incidents per 1,000 population
Percentage of all expedited/remand / full files in time limit which are fully satisfactory or sufficient to proceed
Percentage of full youth files to Crown Prosecution Service both within time which are satisfactory or sufficient to proceed
Percentage of expedited/remand youth files which are fully satisfactory or sufficient to proceed
Number of road traffic collisions involving death or serious injury per 1,000 population

Quality

Percentage of incidents requiring immediate response within target / urban target / rural target
Percentage of 999 calls answered within local target response time
Percentage of persons arrested for notifiable offences who where charged/reported for summons or cautioned
Percentage of persons referred to drug treatment programs

Fair access

Number of PACE stop/searches of white persons per 1,000 population
Percentage of PACE stop/searches of white persons leading to arrest / minority ethnic persons per 1,000 minority ethnic pop.
Percentage of PACE stop/searches of minority ethnic persons leading to arrest
Number of substantiated complaints under PACE 96 detainees per 10,000 people detained
Percentage of racist incidents where further investigative action is taken
Percentage of recorded racially aggravated crimes detected
Percentage of domestic violence incidents (with power or arrest) where arrest was made relating to the incident
Percentage of victims of domestic violence who where victims in the previous 12 months
Percentage of domestic burglaries where the property had been burgled in the previous 12 months

Note: For ease of presentation, multiple, similar performance measures are grouped together and separated by a slash.

Performance measures

There is a long list of Best Value Performance Indicators. The indicators are grouped in five clusters: corporate health, crime, service delivery outcome, quality, and fair access. Table 9.1 provides an overview.

The use of performance measures

Compliance with the performance measures primarily implied reporting back to the Home Office, Police Inspection and Audit Commission. To date it has not been possible to pull this information together to get a clear picture of the overall performance of a force relative to other forces. There is no link between budgets and performance.

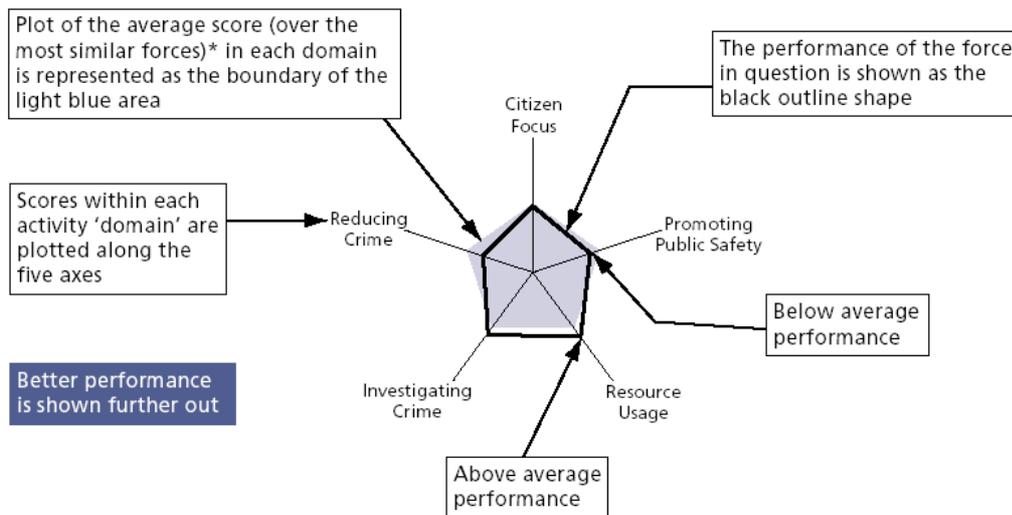
Plans for the future

There are two major policy initiatives. First, the Home Office is developing a method that turns the long list of performance indicators into a meaningful picture of police performance (the so-called 'Policing Performance Assessment Framework'). Six dimensions of police performance are identified: citizen focus, helping the public; reducing crime; investigating crime; promoting public safety and resource usage. Based on social and geographical characteristics, each force is compared to a set of two to nine 'most similar forces'.⁴⁵ Performance monitors will be used to provide a quick and visual representation of performance in these areas. In February 2003, interim indicators have been published based on 2001/02 data (see Home Office, 2003). Figure 9.1 shows an example of a performance monitor.

A second initiative is to create a link between police activities and their costs by 2003/04. Nationally, a system of Activity Based Costing is being introduced. The idea is to use this information to benchmark the forces on their 'value for money'. Based on the outcomes, explicit incentives may be introduced. In the National Policing Plan it says: "The funding regime will need to respond to these developments. The scope for best performing police authorities/forces gaining greater freedom to manage their finances will also need to be examined." (Home Office, 2002, p. 24).

⁴⁵ It is interesting to note that just because one force is 'most similar' to a second force, this does not mean that the second force is most similar to the first force: the second force may actually be closer in character to a third force. For example, one of the 'most similar' forces to a metropolitan force might be an urban force rather than a rural constabulary. However, the urban force might be 'most similar' to other urban forces rather than the metropolitan force (Home Office, 2003, p. 13).

Figure 9.1 Example of a performance monitor



Source: Home Office (2003).

9.3 Policies in Australia

Just as in the last section, we describe the organization of the efforts to measure police performance, the actual performance measures and the way in which the measures are being used.

Organization of the efforts

Australia has a long tradition of comparing the performance of police services between states. The benchmarking efforts are conducted by the Steering Committee for the Review of Commonwealth and State Service Provision, operating under the auspices of the Council of Australian Governments. The Committee comprises senior representatives from Commonwealth, State and Territory governments, and a representative from local government. The two main tasks of the Committee are to develop national performance indicators for government services and to analyse service provision reforms. The services covered by the Committee span education, health, justice, community services, emergency management and housing. A Police Working Group conducts the work on measurement of police performance - in coordination with the Police Practitioners Group.

Performance measures

In January 2003, the Committee published the eighth annual 'Report on Government Services' (see SCRCSSP, 2003, Chapter 5, national policing functions are not included in the analysis).

Over many years, the Committee has worked with the states to improve the quality of performance data and measures. Typically, when there are concerns regarding data comparability or certain performance measures, the Committee initiates a study.⁴⁶ The performance indicators included in the Report measure how well the state police forces meet the four key objectives, given their operating environment. The indicators also measure whether services are provided in an equitable and efficient manner. Table 9.2 provides an overview of the performance indicators for each of the four key objectives for the police.

A new approach to efficiency measurement

The Police Working Group aims to improve current measures of efficiency (ibid., p. 5.58). Typically, these measures are in the form of expenditures for an activity. However, in practice it is hard to attribute costs to one activity. For example, police response to a call for service will not only deal with the incident at hand, but may increase police visibility and, therefore, provide public reassurance. Therefore, a new approach to efficiency measurement is to identify problems of prime importance and the activities required to redress the problems. Measures can then be made of the time/cost of activities and of the actions resulting from those activities.

The use of performance measures

The authors do not offer any possible explanations for differences in measured performance between the states (for example, why are Northern Territory's police expenditures per person double the national average?). It is up to the critical evaluators to ask demand explanations from the police. Clearly, lacking a method of assessing relative performance based on these data, no direct financial incentives are based on these measures.

⁴⁶ Examples of studies that provide input to the police section of the Report on Government Services are: 'Linking inputs and outputs: activity measurement by police services' (1999) and 'Asset measurement in the costing of government services' (2001). See www.pc.gov.au/gsp/

Table 9.2 Report on Government Services Performance Indicators

Key objective (supporting activities)	Indicators
Quality, integrity, access and equity considerations relevant to all key objectives	Percentage of people (very) satisfied with police services (in their most recent contact) Percentage of people who (strongly) agreed the police perform their job professionally Percentage of people who (strongly) agreed that most police are honest Complaints per 100,000 people and per 100 sworn police staff Indigenous staff as proportion of indigenous population aged 20-64 years Percent female police staff
To allow people to undertake their lawful pursuits confidently and safely (emergency assistance; crime prevention activities; community support programs)	Proportion of people who felt (very) safe at home (alone during the day and alone after dark) and in public places (walking or jogging locally after dark, travelling on public transport after dark) Proportion of people who consider the identified issues to be a major problem or somewhat of a problem: housebreaking, drugs, vehicle theft, driver behaviour, physical assault, graffiti/vandalism, family violence, louts/gangs, drunken behaviour and sexual assault Recorded crimes against property and the person and crime victimization Expenditure per person on community safety and support
To bring to justice those people responsible for committing an offence (crime investigations)	Proportion of investigations finalized within 30 days and proportion of finalized investigations leading to offender being proceeded against within 30 days, both for crime against the person and crime against property Expenditure per person on crime investigation
To promote safer behaviour on roads (road safety and traffic management)	Reported use of seat belts, speeding and driving under influence Road fatalities per 100,000 registered vehicles Land transport hospitalizations per registered vehicle Cost per fatal or serious injury or collision Perception of road safety problems in the neighbourhood Expenditure on road safety and traffic management per person and per registered vehicle
To support the judicial process to achieve efficient and effective court case management and judicial processing, while providing safe custody for alleged offenders, and ensuring fair and equitable treatment of both victims and alleged offenders (services to the judicial process)	Death in police custody and custody related operations Proportion of guilty findings in lower and higher court cases Juvenile diversions as a proportion of juvenile offenders Expenditure per person on services to the judicial process Real costs awarded against the police in criminal actions

Source: SCRCSSP (2003).

9.4 What can we learn from these countries?

As stated in the introduction to this chapter, possible lessons concern ideas for policy rather than hard evidence on what works and what doesn't. Based on this overview, we identified the following features that could be of interest to Dutch policy makers.

Involvement of experts

Interestingly, in both countries the policies have been guided by a group of experts. The United Kingdom had its Public Services Productivity Panel (see PSP, 2000) and Australia its Steering Committee for the Review of Commonwealth and State Service Provision (see SCRCSSP, 2002). The role of experts in these two countries stands in contrast with the Dutch strategy. We are not aware of any direct involvement of external experts in the design of the performance contracts for the police forces.

Incentives are based on benchmarking

The requirement to explain differences in performance relative to other forces - rather than the requirement to meet a number of force-specific targets - is used as source of incentives. Both in Australia and in the United Kingdom, critical evaluation of *relative* performance is seen as the key to improve accountability and to strengthen incentives. Obviously, the policy makers realize that many of the performance measures are affected by external factors that may differ between forces. Variation in policing practice is not the only possible reason behind variations. However, as stressed by John Denham, Minister of State for Crime Reduction, Policing and Community Safety, 'variations in policing performance that are highlighted by the comparisons should prompt questions as to why there is such variation' (Home Office, 2003, p. 3). The public, and of course closely involved representatives and government officials, are given the role of critical evaluators who should motivate the forces to improve their performance.

No financial rewards

It is interesting to see that two countries that are way ahead of the Netherlands in quality and organization of police performance data shy away from the use of financial rewards. Apart from the example of asset seizure and forfeiture discussed in Chapter 3, we are not aware of any other countries that use or have used financial incentives for the police. Again, critical questions based on comparisons between forces are the source of incentives, not financial rewards for meeting force-specific targets. Indeed, the United Kingdom is investigating ways of introducing a link between performance and budgets. But they only do so after many years of improving the reliability of police data and research into ways of meaningful comparisons between forces. Moreover, the outcome of these explorations into the use of financial incentives is uncertain (in the next chapter, we provide some reasons for not using financial incentives in the case of

policing). The wish of Dutch policy makers to introduce financial incentives right away does not seem to be based on superior data (see Section 8.3) or a superior way of comparing forces' performance (see Section 8.4).

10 Conclusions

10.1 Aim and structure

Are the problems identified in the second step of the analysis just growing pains of a new policy instrument? Or are there more fundamental problems with performance contracts for the police forces? What are options to proceed - and can we use any of the lessons from policies pursued elsewhere? All these questions address the value of performance contracts for policing, the issue that we will address in this last chapter. Before going into this issue, we first reiterate what goals performance contracts might serve.

10.2 What goals might performance contracts serve?

By holding out the prospect of financial rewards for meeting a set of pre-determined targets, the government wishes to improve the accountability of the regional police forces and strengthen their incentives to work in an efficient, effective, fair and decent manner.⁴⁷

The scope for improvement

The scarce empirical evidence on police performance indicates that there is scope for improvement. Data from the prosecution council shows that there are great differences in performance between the forces. This finding indicates that the forces operate as 'islands' rather than as 'partners in crime' who share best practices. Moreover, no upward trend in police performance can be discerned over the last ten years. To the contrary, the steady increase in the number of police officers has been accompanied by a dramatic decline in efficiency. These findings are consistent with the idea that the police forces do not systematically improve their policing strategies through carefully evaluated experiments. Unfortunately, many of the rich sources of administrative data of the forces have not been opened up yet. Consequently, little is publicly known about the way police budgets have been spent or the results of police work. It is unclear what is behind the decline in measured efficiency. Is the police unable to cope with changes in crime patterns and disorderly behaviour? Or are police officers choked by an ever-increasing administrative burden and a lack of capacity upstream in the criminal justice system? The lack of interest in empirical studies into police efficiency and effectiveness alone is an indication of the low importance attached to understanding and improving police performance.

⁴⁷ When we use the term government, we mean the national government. In police issues, the Interior Department often acts as representative of the national government. Therefore, we sometimes explicitly refer to the Interior Department's role as principal of the regional police forces (see Section 1.4).

The Interior Departments' challenge

The Interior Department's challenge is to find a policy instrument that improves the forces' accountability and incentives, but does not interfere with the discretion of the forces. The great degree of discretion of the forces is based on their superior expertise gained through training and daily experience. The forces, not the policy makers at the Department, know 'where policing is all about' and 'what is really going on in the regions'.

Performance contracts as seemingly attractive policy instrument

At first sight, entering into performance contracts with the forces seems to be an attractive way of improving accountability and incentives. Holding the forces accountable to results acknowledges the information asymmetry between the Interior Department and the forces regarding the best way to improve performance. Performance contracts combine the two policy instruments of the Department: exercising oversight and budgeting. The rewards for meeting a set of targets provide financial incentives that support the forces' drive to improve. The rewards compensate the forces for the time and effort spent on evaluating operations and motivating employees to share and implement best practices. Moreover, the rewards acknowledge the efforts of police officers and their managers who have to overcome (bureaucratic) barriers to change and the continuous, strong pressure to respond to the issues of the day.

10.3 Are the proposed performance contracts likely to be effective?

Putting desirable improvements in performance into a contract poses some serious problems in the case of policing. Considering the way in which police performance is measured and rewarded, the proposed performance contracts are not likely to be uniformly beneficial to the realization of the government's objectives for the police forces. We base this conclusion on the following findings.

Output measures invite gaming

The output measures - the number of fines and the number of charges forwarded to the prosecution council - invite the forces to game the performance contract. Gaming refers to narrowly focussing on 'meeting the numbers' at the cost of other dimensions of police work not included in the contract. When requiring the forces to meet one-dimensional output targets, the wrong behaviour is also rewarded.⁴⁸ The forces are rewarded for substituting away from tasks

⁴⁸ The prosecution council tries to put a brake on undesirable responses in the area of crime fighting by using its own performance measures, prescribing desirable behaviour and seconding employees to police stations ('secretaris op locatie'). The temptation to game the evaluation remains, however. The (financial) incentives are based on the Interior Department's performance measures after all. Other performance measures included in the contract do not provide a sufficient check on gaming either.

that do not affect the performance measures to tasks that do (at the possible cost of ‘pro-active’ crime prevention). Within activities that affect the measures, the performance contracts also pay off for focussing on the easiest cases (frivolous fining and picking the easiest criminal charges). Finally, the forces are rewarded for lowering quality of output (forwarding charges with questionable evidence to the prosecution council). If the forces are expected to react to the (financial) incentives, then performance contracts that pay off for behaviour other than what the government is seeking are not likely to be effective.

Output targets conflict with the forces’ discretion

By putting the desired number of fines and criminal charges in a contract, they are made independent of changes in the nature, frequency and location of criminal and disorderly behaviour. Therefore, the demands of performance contracts and reality may well start to deviate. Such a gap between reality and targets limits the responsiveness of regional police work to changes in the environment. It is in the government’s interest to provide the forces with the discretion to decide when to fine someone or when to start a criminal investigation (although the forces can use decision tools provided by third parties such as the prosecution council). After all, the police forces have the information to respond to changing regional circumstances.

Financial rewards are a further invitation to gaming

Financial rewards require translating multiple dimensions of police work into one binary outcome. Allocating the reward is a binary decision after all. Either the force meets the conditions for reward or it does not. In contrast, police work has many, non-discrete outcomes such as reassuring citizens and providing assistance at traffic accidents. Weights need to be attached to all of the different outcomes. It is hard to replace subjective judgement with mechanical rules, however. When the weights included in such ‘decision rules’ are picked wrong, undesirable behaviour is rewarded. The forces can be expected to look for the target that can be achieved with the least effort and the greatest reward after all. At the time of our writing, it is unclear how the Department will address this challenge. Given its serious shortcomings, the now-abolished system of performance rewards does not provide much guidance.

Financial rewards conflict with the lack of consumer choice

By not increasing the budgets of forces that did not sufficiently improve their performance, citizens pay twice for poor police management within their locality.⁴⁹ In contrast with schools and hospitals, citizens cannot switch to a better performing force other than by moving somewhere else.

⁴⁹ For the moment, we assume that meeting the targets included in the performance contracts is equivalent to improving police performance in the perception of citizens.

Performance targets that are off undermine incentives

Back-of-the-envelope calculations indicate that the targets for the number fines and charges seem to be too little ambitious - certainly given the future increase in the number of police officers. The goals for customer satisfaction may also be little ambitious since they are based on historical performance ('perform at least as good as you did in the past'). In the absence of any firm knowledge about the means by which the targets could be achieved, each target included in the performance contract is likely to be either out of reach or too little ambitious. If they are set too low, the force may stop their efforts as soon as the target is reached ('satisficing behaviour'). In the case that targets are too ambitious, the forces also have little incentive to improve performance. Therefore, the performance contracts may not provide much incentives to improve performance.

Wish for flexibility at the national level may undermine power of incentives

The government's wish for flexibility in changing policing priorities and the terms of contract may undermine the performance contracts' incentives. In practice, the government frequently places new demands on the police. Through such interventions the government becomes a 'co-producer' who shares responsibility for the results. Therefore, such interventions may lower the incentives to meet the performance targets. Furthermore, negative experiences with some of the performance measures may prompt the government to change the terms of contract. In July 2003 the current state of the art of measuring and rewarding police performance will be frozen into performance contracts. Given the relative inexperience with this new policy instrument - and the above-mentioned shortcomings in contract design - changes in the terms of contracts may become desirable. Changes in the terms of contract will undermine the credibility of the Department as principal, however, and therefore weaken incentives.

To conclude, a closer look at the proposed performance contracts learns that they may not be entirely effective in advancing the government's agenda of improving police accountability and incentives. Strengthening incentives requires more of target setting than may be possible. Moreover, some performance targets can also be met by strategies that are not in line with the government's objectives. Finally, the performance contracts conflict with the desired flexibility at the regional and national level and the lack of consumer choice. Therefore, it should be interesting to compare performance contracts with other approaches used elsewhere.

10.4 What are options for policy?

Some of the above-mentioned problems could be alleviated by making a number of changes to the performance contracts. However, serious drawbacks remain with the contractual approach. Based on our brief look into policies pursued in Australia and the United Kingdom, we outline

an alternative approach that may provide a better fit with the nature of police work and the desired discretion of the forces.

Some ways of improving upon the proposed performance contracts

Better data, other performance measures and a redirection of financial incentives to police management could alleviate some of the problems with the proposed performance contracts:

- *Search for better performance indicators.* Better measures could limit the side-effects and lower the conflict with flexibility at the regional level. Of all measures included in the contract, the number of fines and charges most strongly invite the forces to just 'meet the numbers'. They also limit the discretion of the forces most strongly. An option is not to include the number of fines in the contract. The charges can be weighted with their respective penalty. Then the forces no longer have the incentive to focus on the easiest cases.
- *Tie financial incentives to police managers' remuneration rather than the forces' budgets.* That way, customers of police services will no longer pay twice for poor police performance.
- *Gather better data to improve target setting.* Better data on how the forces spend their resources and what outputs are produced are necessary to make a better informed guess on what the targets should be. To that end, the rich sources of administrative data of the forces would need to be opened up.

Serious drawbacks remain

The major difficulty with performance contracts is the need to fix the desired improvements in police performance *in advance*. The contractual approach is weak on unexpected behavioural responses to performance targets and conflict with the necessary flexibility at the regional and national level. Moreover, thorough assessment of a force's strategy is impaired since police performance is reduced to a small number of quantitative performance measures that everyone can understand. The rules that need to be established to allocate the financial rewards only worsen these problems. Clearly, these are problems that are inherent to the contractual approach. They cannot be fixed with better performance measures, targets and the like.

Looking for an alternative approach

Based on our look at policies across the borders in Chapter 9, we distinguish two approaches to achieving better police accountability and incentives. One approach is to contract desirable improvements in the performance of the police forces in advance. The Dutch government largely follows this approach. The other approach is to let customers and peers critically evaluate achieved results - with an eye on meeting future challenges. The latter approach is more in line with policies pursued in Australia and the United Kingdom. This approach has also been vigorously put into practice by William Bratton, former New York City Police Commissioner and

currently Chief of the Los Angeles Police Department (see Kelling, 1995 for an informative description). As a conclusion to this study, we will compare and contrast the two approaches and discuss their likely effectiveness in achieving the government’s objectives for the police forces. We provide an overview of the key characteristics of the approaches in Table 10.1.

	‘Contracting desired changes’	‘Critical evaluation of achievements’
Practical form	Performance contracts with financial rewards for meeting targets	Customer satisfaction surveys (citizens, public prosecutors, businesses), managed peer review
Source of incentives	The requirement to meet pre-determined, force-specific performance targets	The requirement to explain achieved results relative to other forces to customers and peers
Type of reward	Financial bonus for meeting targets, good publicity when accomplishments are publicly announced	Remuneration for accountable managers, acknowledgment by evaluators, good publicity when accomplishments are publicly announced,
Role of national government	Fixing desired changes in a contract	Organizing peer reviews and customer satisfaction surveys
Necessary ingredients	‘Gaming proof’ performance measures, reliable, non-manipulated performance data, information to set targets and financial rewards	Unbiased evaluators, reliable and comparable information on regional situation, policing strategies and achieved results

How could a policy based on a ‘critical evaluation of achievements’ look like?

Based on the foregoing analysis of the Dutch initiatives, we have some idea of the ‘contracting desired changes’-approach. How could a policy based on ‘critical evaluations of achievements’ look like? The basic premise of this approach is that *the interaction with customers and other forces - structured by the government - drives the motivation to improve performance*. Customer surveys and managed peer reviews could be organized to facilitate this interaction.

Customer surveys

Perceptions of customers can be obtained through surveys. Citizens are already being surveyed through the Citizen Police Monitor (‘Politie Monitor Bevolking’). Business can be surveyed in line with the questions in Section 4 of the ‘Monitor Bedrijven en Instellingen’, see Visser et al. (2002a, p. 72). Public prosecutors can be surveyed following the ‘Bevoegd Gezag Monitor’, an initiative of police force Brabant Zuid Oost. The results of the customer surveys can then be published in line with the government’s Safety Plan:

“We will also work towards publicly disseminated comparisons between the performance of the regional police forces” Justitie and BZK (2002, p. 82).⁵⁰.

Subjective assessment allows a more complete picture of performance to be attained, not possible with objective measures. Each of these groups of customers can provide useful feedback on the specific part of police services that they are able to assess. Customers are likely to give their honest opinion about the police in their region. Therefore, the assessment will not be tainted by hidden agendas. Most importantly, subjective performance measures are relatively ‘gaming proof’ compared to output measures such as the number of criminal charges. After all, customers will be on the alert for quality of police work. Moreover, there does not seem to be much scope for inefficient ‘flattering’ of customers (think of the salesman who gives too much price brakes ‘to please customers’). For example, businesses state that the determinants of their satisfaction with police services include speed of recording a crime, the capability to solve the problem at hand, and courtesy in contacts with the police (Visser et al., 2002b, p. 64). Efforts that are aimed at these factors are perfectly in line with the government’s objective of efficient, effective, fair and decent police work. Moreover, these are all things the police can do something about. There is a clear link between effort and measured performance.

Managed peer review

In addition to customer satisfaction surveys, managed peer reviews are a necessary element of the assessment. Customers do not look at the interrelations between the complete set of police activities. Moreover, they may not take into account all of the dimensions of police performance that are relevant to the government. For instance, customers may not critically review the efficiency of the police services delivered. Therefore, it is the role of experts to critically assess the overall strategy of the police force. Peers play a prominent role here, since there are no better experts than peers. To ensure unbiased evaluations, the peer review could be managed by an external expert. The managed peer review could probably be built upon the existing peer reviews that are part of the system of quality management (‘Kwaliteitsstelsel politie’, see Chapter 3).

A force’s management can be demanded to explain noticeable differences with other forces and also their plans to address identified problems and new challenges. For example: ‘Your commercial robberies are high compared to the Haaglanden region. How many of the robberies are kids stealing cupcakes? How many are guys walking in with Uzis?’ And: ‘Why didn’t you follow the strategy used in Rotterdam Delfshaven?’ etc. (free after Kelling, 1995). To be able to answer these questions, a force’s management needs to have a coherent picture of the problems

⁵⁰ Translated from: “Tevens wordt het periodiek publiekelijk bekend maken en vergelijken van resultaten in de verschillende politieregio’s gerealiseerd.”

in their region, their strategy to tackle these problems, and the results. Clearly, such a picture needs to be based on close relations with the rest of the force ('How did you guys tackle that burglary problem?') and sound data. Recent reports by the General Accounting Office show that major improvements are possible in the management information systems that are necessary to support this approach (Algemene Rekenkamer, 2003a, 2003b).

Conclusion: contrasting the two approaches

As Table 10.1 already indicated, the two approaches are fundamentally different. In the current proposals, the financial rewards for meeting a set of pre-determined targets is seen as the way to improve police accountability and incentives. Following a policy of 'critical evaluation of achievements', the source of incentives is the requirement to answer critical questions about relative performance from customers and peers. They possess the information to conduct a thorough assessment of police performance - *and to provide ideas about ways to improve*. After all, customers are being served by the police and peers are in the business of policing. Who else is in a better position to evaluate the forces? Appreciation from customers and peers forms the reward for achievements rather than financial bonuses from the Department.

This approach acknowledges the discretion of the police at the regional level. The role of the national government is not to put down desired changes in a contract but to organize the interaction between a force's management and their customers and peers. By doing so the forces are made responsible for presenting a coherent picture of the problems in their region, their strategy to tackle these problems, and the results. The ability of police managers to be accountable to - and to learn from - customers and peers can be an input into decisions about promotion and remuneration. Their performance is assessed based on their professionalism as police manager rather than on meeting a number of quantitative performance targets.

Therefore, it is worthwhile to look into this alternative approach towards improving accountability and incentives. The government may be able to better realize its objectives for the regional police forces. The help of experts and experiences in countries such as the United Kingdom, Australia and also the United States can be useful in designing and implementing a policy that supports a culture within the police of systematically following, evaluating and improving performance.

Abstract

This year, the government will enter into performance contracts with each of the 25 regional police forces. The performance contracts establish a direct link between meeting a number of quantitative performance targets and financial incentives. A major improvement in police performance is necessary to meet the objective of 20 to 25 percent less criminal and disorderly behaviour by 2006.

A closer look at the performance contracts learns that they may not be the most appropriate policy instrument to achieve this objective. The nature of police work does not allow for advance planning of outputs. The police consist of professionals who need a high degree of discretion to do their work. The targets invite adverse behavioural effects. Management could become focussed on 'meeting the numbers' rather than on delivering results. Because of the wide variety in police tasks and the low measurability of quality, there is a wide gap between performance measures and results. The financial incentives make it worse, by forcing a yes/no decision based on weighing multiple, non discrete performance measures. Moreover, the targets are likely to be off since the government does not have the information to set them at the right level. Less financial resources for poorly performing forces also adversely affect citizens. They cannot choose between providers of police services as in the case of hospitals or schools.

Experiences in Australia and the United Kingdom suggest an alternative approach. They focus on benchmarking of police forces without direct financial incentives. Both countries have invested many years in improving the quality and comparability of police data as well as methods for fair comparisons between forces. Based on these comparisons, police forces are held accountable. Consequently, the police are being forced to develop a clear picture of the effects of their approach in terms of the region's specific problems. The Dutch government could follow a similar approach. A system of peer review and customer satisfaction surveys can be instrumental in assessing a force's performance and in providing ideas for improvement. Critical assessment of performance data by knowledgeable people is a necessary ingredient to a policy of holding the police accountable to results. It stimulates a culture of experimenting, data collection and analysis, and singling out and sharing best practices. Such a change is necessary to bring about the desired improvement in police performance.

References

Algemene Rekenkamer, 2001, *Handhaving door Rijksinspecties*, TK 28 271, Sdu Uitgevers, The Hague.

Algemene Rekenkamer, 2003a, *Zicht op taakuitvoering politie*, TK 28 791, Sdu Uitgevers, The Hague.

Algemene Rekenkamer, 2003b, *Uitwisseling opsporings- en terrorisme-informatie*, TK 28 845, Sdu Uitgevers, The Hague.

Arnott, Jenny and Kaite Emmerson, 2001, *In Sickness and in Health: Reducing sickness absence in the police service*, Home Office, Police Research Series Paper 143, London.

Baker, George, 2002, Distortion and Risk in Optimal Incentive Contracts, *Journal of Human Resources*, 37 (4), pp. 728-751.

Blumstein, Alfred, 1999, Measuring what matters in policing, *Proceedings from the Policing Research Institute meetings*, US Department of Justice, pp. 5-9.

Boom, Wilco, 2002, Politie moet weer gaan speuren, *Algemeen Dagblad*, May 14.

Bosker, Okko Jan, 1997, *Snelrecht - de generaal en speciaal preventieve effecten van sneller straffen*, Thesis, Rijksuniversiteit Groningen, Groningen.

BZK (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2001, *Landelijke politiebrieven 2002*, The Hague.

BZK (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2002a, *Kerngegevens Nederlandse Politie 2001*, The Hague.

BZK (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2002b, *Politiesterkte eerste helft dit jaar gegroeid*, News item on www.minbzk.nl, The Hague.

BZK (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2002c, *Budget 2003*, The Hague.

- BZK (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2002d, *Brief aan de Tweede Kamer der Staten Generaal*, The Hague.
- BZK and Justitie (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Ministerie van Justitie), 1998, *Beleidsplan Nederlandse Politie 1999-2002*, The Hague.
- BZK and Justitie (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, Ministerie van Justitie), 2003, *Landelijk Kader Nederlandse Politie 2003-2006*, The Hague.
- Carrington, Roger, Nara Puthuchery, Deirdre Rose and Suthathip Yaisawarn, 1997, Performance measurement in government service provision: the case of police services in New South Wales, *Journal of Productivity Analysis*, 8, pp. 415-430.
- Cornet, Maarten and Jeroen van de Ven, 2003, Prikkels rond TNO, *mimeo*, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.
- De Mulder, Richard, 2002, 'Veiligheid' bepaalt verkiezingsuitslag, *Algemeen Dagblad*, April 6.
- Dijk, Tom van, Sander Flight and Erik Oppenhuis, 2000, *Voor het beleid achter de cijfers. De uitkomsten van de GSB-monitor veiligheid en leefbaarheid nader geanalyseerd*, Intomart, Hilversum.
- Drake, Leigh and Richard Simper, 2001, *The Economic Modeling of Policing and the Measurement of Efficiency*, Economic Research Paper No. 01/04, Loughborough University, Loughborough.
- Drake, Leigh and Richard Simper, 2002, English and Welsh Police Force Efficiency: Demonstration Project Peer Review Results, *mimeo*, Loughborough University, Loughborough.
- Dunn, Kyla, 2000, *Reining in forfeiture: common sense reform in the war on drugs*, PBS Frontline 'Drug Wars' series, Washington D.C.
- Dupont, Benoit, 2001, The new governance of Australian police services, *mimeo*, Charles Sturt University, Goulburn.
- Eisenberger, Robert and Judy Cameron, 1996, Detrimental Effects of Reward. Reality or Myth?, *American Psychologist*, 51 (11), pp. 1153-1166.
- FitzGerald, Marian, Mike Hough, Ian Joseph and Tariq Qureshi, 2002, *Policing for London*, Willan Publishing.

- FRP (Federatie van Regionale Politievakorganisaties), 2000, Audit Rapportage Politie Gelderland Zuid, *mimeo*.
- Goldstein, Herman, 1963, Police Discretion: The Ideal Versus the Real, *Public Administration Review*, 23.
- Hart, Oliver, Andrei Shleifer and Robert W. Vishny, 1997, The proper scope of government: theory and application to prisons, *Quarterly Journal of Economics*, 112 (4), pp. 1127-1161.
- HMIC (Her Majesty's Inspectorate of Constabulary), 1998, *What Price Policing? A Study of Efficiency and Value for Money in the Police Service*, London.
- Holmstrom, Bengt, 1979, Moral hazard and observability, *Bell Journal of Economics*, 10, pp. 74-91.
- Holmstrom, Bengt and Paul Milgrom, 1991, Multitask principal-agent analysis: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics & Organization*, 7 (Special issue), pp. 24-52.
- Home Office, 2002, *The National Policing Plan 2003-2006*, London.
- Home Office, 2003, *Police Performance Monitoring 2001/02*, London.
- Justitie and BZK (Ministerie van Justitie, Ministerie van Binnenlandse Zaken en Koninkrijksrelaties), 2002, *Naar een veiliger samenleving*, The Hague.
- KBB (Korpsbeheerdersberaad), 2001, *Prestatiebeloning bij de Nederlandse Politie. Een aantrekkelijke buidel*, The Hague.
- KBB (Korpsbeheerdersberaad), 2002, *Prestatiebeloning - Eindadvies inzake verdeling budget 2002*, The Hague.
- Kelling, George L., 1995, How to Run a Police Department, *City Journal*, 5 (4), pp. 34-45.
- Kleemans, E.R., M.E.I. Brienens and H.G. van de Bunt, 2002, *Georganiseerde criminaliteit in Nederland*, Tweede rapportage op basis van de WODC-monitor, No 198, Ministerie van Justitie, The Hague.

Koning, Pierre W.C. and Anja Deelen, 2003, Prikkels voor het UWV, *mimeo*, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

Kuhry, Bob and Ab van der Torre, 2001, *Quartaire Sector*, Chapter 9 in: Sociaal Cultureel Planbureau, De sociale staat van Nederland, The Hague.

McCrary, Justin, 2002, Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment, *American Economic Review*, 92 (4), pp. 1236-1243.

Muller, E.R., 2002, Policing and accountability in the Netherlands: a happy marriage or a stressful relationship?, *Policing and Society*, 14 (4), pp. 249-258.

Nationale Ombudsman, 2001, *Jaarverslag 2001*, The Hague.

New Zealand Police, 2002, *Judicial Satisfaction Survey 2001-2002*, Office of the Commissioner, Wellington.

OECD, 1999, *Performance contracting, Lessons from performance contracting case studies and a framework for public sector performance contracting*, Public Management Committee, PUMA/PAC(99)2, Paris.

Openbaar Ministerie, 2002, *Politiebenchmark, mimeo*.

Openbaar Ministerie, 2003, *Aanwijzing voor de opsporing*, College van Procureurs-Generaal, The Hague.

Pate, A., M. Wycoff, W. Skogan and L. Sherman, 1989, *Reducing fear of crime: a summary report*, Police Foundation, Washington.

PKP (Project Kwantificering Politiewerk), 1988, *Eenentwintig politiekorpsen met gelijke maat gemeten*, Andersson Elffers Felix, Utrecht.

Polinsky, A. Mitchell and Steven Shavell, 2000, The Economic Theory of Public Enforcement of Law, *Journal of Economic Literature*, 38, March, pp. 45-76.

Prendergast, Canice, 1999, The provision of incentives in firms, *Journal of Economic Literature*, 37 (1), pp. 7-63.

- Prendergast, Canice, 2001, *Selection and oversight in the public sector, with the Los Angeles Police Department as an example*, NBER Working Paper 8864, Cambridge (MA).
- Prendergast, Canice, 2002, Uncertainty and Incentives, *Journal of Labor Economics*, 20 (2), p. S115-S137.
- PSP (Public Service Productivity Panel), 2000, *Improving police performance. A new approach to measuring police efficiency*, HM Treasury, London.
- Rovers, Ben, 1999, Het aantal opgehelderde misdrijven en gehoorde verdachten neemt al jaren gestaag af, *Tijdschrift voor Politie*, 11, November, pp. 10-14.
- Schick, Allen, 1996, *The Spirit of Reform: Managing the New Zealand State Sector in a Time of Change, A report prepared for the State Services Commission and the Treasury*, New Zealand, Wellington.
- Schoof, Rob, 2002, 'Roep om meer veiligheid is in hoge mate symbolisch' - Onderzoeker Frits Vlek over de meetbaarheid van politiewerk, *NRC Handelsblad*, October 10.
- SCP (Sociaal Cultureel Planbureau), 2001, *De sociale staat van Nederland*, The Hague.
- SCP (Sociaal Cultureel Planbureau), 2002, *Sociaal en Cultureel Rapport 2002*, The Hague.
- SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision), 1999, *Linking inputs and outputs: activity measurement by police services*, Melbourne.
- SCRCSSP (Steering Committee for the Review of Commonwealth/State Service Provision), 2003, *Report on Government Services 2003*, Melbourne.
- Shirley, Mary M. and Lixin Colin Xu, 2001, Empirical effects of performance contracts: evidence from China, *Journal of Law and Economic Organization*, 17 (1), pp. 168-200.
- Skolnik, Sam, 2001, Critics target drug raid seizures, *Seattle Post-Intelligencer*, December 13.
- Smit, P.R., 2000, Nederland in internationaal perspectief, Chapter 8 in: F.W.M. Huls, M.M. Schreuders, M.H. ter Horst-van Breukelen and F.P. van Tulder, *Criminaliteit en rechtshandhaving 2000*, Report No. 189, Department of Justice, The Hague.

Staffordshire Police, 2001, *Deputy refutes third worst force allegation*. Response to 'The Observer', Staffordshire, News Item, December 5. www.staffordshire.police.uk/news061.htm

Tak, P.J.P. and J.P.S. Fiselier, 2002, *Duitsland-Nederland; en de afdoening van strafzaken*, Wolf Legal Productions.

Terpstra, Jan, 2002, *Sturing van politie en politiewerk - Een verkennend onderzoek tegen de achtergrond van een veranderende sturingscontext and sturingsstijl*, IPIT/Politie en Wetenschap, Enschede.

TK26 345 (Tweede Kamer der Staten-Generaal), 2002, *Vaststelling van de begroting van de uitgaven en de ontvangsten van het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (VII) voor het jaar 2002*, The Hague.

TK 27 834, nr. 2 (Tweede Kamer der Staten-Generaal), 2001, *Nota Criminaliteitsbeheersing*, The Hague.

Torre, E.J. van der and W. Ph. Stol, 2000, *Waardevolle politieverhalen: politie en Marokkaanse jongens*, COT Universiteit Leiden/Nederlandse Politie Academie, Elsevier.

Travis, Jeremy and Michelle Waul, 2002, *Reflections on the Crime Decline: Lessons for the Future?*, *Proceedings from the Urban Institute Crime Decline Forum*, Urban Institute, Washington.

UK Parliament, 1999, *Experiences in New Zealand*, Metropolitan Police Budget [24 March 1999] Column 331, London.

Verhoeven, W.H.J., P.A. van der Hauw, K.L. Bangma, J. de Muijnck, F.M.J. Westhof, 2002, *Ontwikkeling arbeidsproductiviteit in de publieke sector, 2003-2006*, EIM, Zoetermeer.

Vijver, C.D. van der, A.J. Meershoek, D.F. Slobbe, 2001, *Kerntaken van de politie. Een inventarisatie van heersende opvattingen*, IPIT/Politie & Wetenschap, Twente University, Enschede.

Visser, J., R. Fredrikse and E. Hermans, 2002a, *Handboek Monitor Bedrijven en Instellingen 2002*, NIPO, Amsterdam.

- Visser, J., R. Fredrikse and E. Hermans, 2002b, *Slachtofferschap criminaliteit bij bedrijven en instellingen. Monitor Bedrijven en Instellingen: nulmeting onder 5.000 vestigingen*, NIPO, Amsterdam.
- Vollaard, Ben A., 1999, Arbeidsbemiddeling in Australië, *Economisch Statistische Berichten*, 84 (4213), pp. 539.
- Wiebrens, Caspar, 2002a, Raak geschoten, mis geschoten of helemaal niet geschoten? Ervaringen met case screening bij de Nederlandse politie, *mimeo*, Openbaar Ministerie, The Hague.
- Wiebrens, Caspar, 2002b, Celdagen zeggen meer over doelmatigheid politie, *Economisch Statistische Berichten*, 87 (4389), pp. 900-903.
- Wiebrens, Caspar and Sjaak Essers, 1999, Schaf het ophelderingspercentage af, *Tijdschrift voor de Politie*, 61 (1), pp. 27-34.
- Wittebrood, Karin, 2002, *De prestatieparadox van politie en justitie*, Nieujaarspocket 2002, SCP, The Hague.
- Zoomer, Olga, 2001, Gebiedsgebonden politiezorg, *InterDisciplinair*, 12 (1), pp. 22-24.

Annex 1. The measures for the system of performance rewards

Performance measures included in the (now abolished) system of performance rewards

Output and outcome measures

Percentage of survey respondents that feel safe

Willingness to report a crime and actual reported crimes per 1,000 inhabitants

Average score on five survey questions on public space (graffiti, harassment by young people, waste, dog dirt, vandalism of street benches etc. and noise nuisance)

Average score on four survey questions on neighbourhood traffic (road hogs, noise nuisance, accidents)

Number of reported raids on companies/institutions per officer

Number of investigations related to illegal trade, moral law and the environment forwarded to the public prosecutor per officer¹

Number of persons charged in investigations related to illegal trade, moral law and the environment forwarded to the public prosecutor per officer

Subjective performance measures

Average score on five survey questions on police services (police protects citizens, police has contact with people in the neighbourhood, police does a good job in the neighbourhood, and the police has an efficient approach to problems)

Percentage of survey respondents that are (very) satisfied with their last contact with the police

Average score on five survey questions on accessibility of the police

Measures of internal performance

Percentage of persons charged with home / street / physical threat-related offences forwarded to the public prosecutor

Number of arrests divided by the number of traffic fines per officer

Number of charged offenders who are forwarded to the public prosecutor per 100 reported offences

Number of underage suspects referred to HALT divided by the total number of underage persons charged with offences²

Percentage of charges dismissed by the public prosecutor (technical dismissal type 2: inconclusive evidence)

Number of days for processing of reported crimes

Net number of worked days (work days excl. leave, sickness, training) as percentage of gross number of hours of work

Notes: (1) Illegal trade refers to drugs trade, frontier-running, weapons trade and fraud; (2) HALT is aimed at young offenders (12-18 years of age). They get the opportunity to repair or pay for the damage caused by their behaviour. They can do so by doing supervised work. The program is targeted at youth involved in vandalism and small theft.

Annex 2. People consulted

The following people were consulted during the course of the study. Additionally, I had a number of informal conversations with police officers, including Rotterdam-West community beat officer Th. J. van Asperen. They were helpful in either shaping or testing the ideas and recommendations in this report, without in any way being responsible for them.

Name(s)	Institution	Place, Date
Sjoerd van Dommelen, Arjen Gielen	Ministerie van Financiën	Den Haag, September 26, 2002
Cees van der Vijver	Universiteit van Twente	Enschede, October 1, 2002
Claartje In 't Velt, Wouter Stol	Nederlandse Politie Academie	Amsterdam, October 4, 2002
Caspar Wiebrens	Openbaar Ministerie	Den Haag, October 7, 2002
Frits Vlek	Commissie Politie & Wetenschap	Apeldoorn, October 23, 2002
Cas Schreuder	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties	Den Haag, November 7, 2002*
Frank Bovenkerk	Universiteit van Utrecht	Utrecht, November 12, 2002
Rob Merten	Nederlands Politie Instituut	Den Haag, November 25, 2002*
Marlies Burm	Algemene Rekenkamer	Den Haag, November 26, 2002
Marc Jacobs	Politiedistrict Utrecht-Binnensticht	Zeist, November 28, 2002
René Goudriaan	APE Consultancy	Den Haag, February 3, 2003
Benoit Dupont	Université de Montréal, Canada	Den Haag, February 17, 2003*
Andrew Hughes	New South Wales Treasury, Australia	Den Haag, February 28, 2003*
Jan Haagsma	Andersson Elffers Felix	Den Haag, March 7, 2003
Oskar Huurdeman, Esther de Kleuver	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties	Den Haag, March 10, 2003
André de Jong, Kees Wilmer	Ministerie van Financiën	Den Haag, March 21, 2003
Jack Riley	RAND Corporation	Los Angeles, April 18, 2003

Note: (*) Communication by electronic mail.