

Summary

Benchmarking in the prison system

A study on the possibilities of comparing and improving performance

Cause

This study has been requested by the Dutch National Agency of Correctional Institutions (DJI) and the Directorate General of Prevention, Youth and Sanctions (DGPJS). DJI is responsible for the execution of prison sentences and judicial measures, and falls under the policy responsibility of DGPJS. The prison system is DJI's largest sector, confronting it with a growing need for information among, for instance, administrators, politicians and inspectors. They ask an increasing number of questions about the implementation, whether or not prompted by incidents. As a result, the starting point arising from the established relation between DJI and the Ministry of Security and Justice (DJI is a cost-benefit service), of accountability according to main lines, is subject to growing pressure. For this reason, the sector is left with less scope to deal with problems innovatively and to search for ways to do better. Against this background, DJI and DGPJS want to explore the options to improve their grip on the implementation of tasks by the correctional institutions. These tasks relate to the fields of safety and the humane treatment and rehabilitation of inmates. The objective is to find an instrument with which the performance of different institutions can be *compared*, while at the same time, it will also help with *learning, improving, and accountability*. The WODC has been asked to examine the potential of benchmarking as an instrument to make the performance levels more transparent and to stimulate the learning- and improving process, thus strengthening the trust between administrators and executing agencies. The central question of this study is:

How can we draw an integral comparison between the performances of correctional institutions, and in what way should benchmarking take place if we want to stimulate learning, improvement and accountability?

Method

For this study, we have used the literature from different scientific fields, such as public administration and penology. We have also carried out quantitative analyses of empirical figures from the Dutch prison system. In addition to this, we have consulted experts working within the prison system.

Findings

What does benchmarking entail?

Benchmarking came into being in the business world, as a method to enable organisations and parts of organisations to 'copy each other's tricks', to make them more successful in realising their objectives. As an instrument, benchmarking has come to absorb more and more functions, such as the monitoring and comparing of performances, the transference of *good practices*, and providing management information. The comparison of performance and the uncovering of differences are meant to incite organisations to action.

Is benchmarking also suitable for the public sector?

The usefulness of benchmarking for the public sector is not evident. For public organisations, a bad performance does not immediately turn into a threat to their existence, because the free market process affects them less. In the meantime, however, different branches within the public sector have acquired positive experiences with benchmarking. Providing insight into differences in performance incites to action and measures, but the public organisations' administrators should see to it that the instrument is used actively.

Which data can be used to measure achievements?

Experiences from other branches at home and abroad show that measurements regarding rather abstract objectives ('outcome', which for instance relates to public safety and the prevention of recidivism in the case of correctional institutions), are as yet unsuitable to give a picture of the performance of public organisations. The contribution of a public organisation to such phenomena is uncertain and indirect at best, hard to trace and just one of the factors at work. Instead, it seems a better idea to use the direct results of the implementation ('output') and the way in which these results have been achieved ('throughput').

How should achievements be valued?

Benchmarking in the public sector does not need to aim for the highest imaginable achievement on every indicator, as is often the case in the business world. It is not desirable, for example, that the inmates live in luxury; on the other hand, it is regulated by law that inmates must have access to particular services, such as sports facilities, a visiting area, and a library. Thus, a standard needs to be laid down for the desired performance level. In addition to effectiveness, efficiency is important as well. When two organisations both meet the standard, the one with the lowest cost performs the best.

What is the definition of benchmarking in the public sector?

Benchmarking in the public sector is an instrument to be used for the purpose of operational management and accountability, focused on the most important goals of an organisation. It measures and compares the performances of this and other, similar organisations and lays down a standard of achievement by appointing one organisation as the ‘desired level performer’. It connects this organisation with others that deviate from the standard, to exchange and apply ‘good practices’ in order to improve the overall performance.

Which themes should be the subject of the performance measurement and the benchmarking?

The prison system executes the prison sentences and judicial measures imposed by a judge. The judge’s objective in imposing detention is therefore the first thing to go on. The so-called objectives of punishment are repayment, specific prevention and general prevention. Not every one of these objectives of the judge has a direct effect on the implementation by the prison system.

The areas in which correctional institutions do need to perform are:

- the safety of inmates and staff in the prison (internal safety);
- no escapes (external safety);
- human dignity;
- rehabilitation.

From these performance themes arise all kinds of tasks. In this study, we have mapped these out based on relevant Dutch policy documents, legislation and rules.

Measurements are available on the theme of safety. The planning and control cycle and surveys among personnel and inmates, for instance, provide data on safety. With regard to other task areas, nationwide data are available only to a limited extent. For this reason, we have decided to limit our study to the performance theme of safety.

Which limiting factors should be taken into account when carrying out a performance measurement?

The available data on safety do not always clearly show the extent to which a correctional institution controls the safety situation. When carrying out a performance measurement, we assume that the scores can be attributed to the accomplishments of the management and staff, but this is not necessarily true. The organisation’s situation can play a role as well, like for instance the region, the building, or the type of inmates housed by the prison. For this reason, a useful performance comparison takes into account the ‘given circumstances’. In this study, we have called these *restrictive factors*. From the literature, we have extracted an overview of the factors that are important with

regard to the safety in a correctional institution. Subsequently, we have asked experts to assess the extent to which these factors can be influenced.

Restrictive factors regarding the performance theme of safety are:

- characteristics of the inmates, such as age, the number of offences committed earlier, ethnicity, and possible emotional and mental disorders;
- characteristics of the institution, such as the use of one cell by more than one individual, the size of the institution, the architectural set-up of the cell complex, and the open or closed nature of the regime.

Influenceable factors are:

- the composition of the prison population per living unit;
- the composition of the work force;
- the circumstances of detention;
- security measures;
- aspects of staff management, such as styles of leadership, communication between staff members, and the ways in which inmates are treated.

Which factors are important for safety?

For this study, we have carried out quantitative analyses to check whether the restrictive factors and the influenceable factors are correlated to the scores of correctional institutions on safety indicators. These analyses show that for most indicators, the scores correlate to both restrictive factors and influenceable factors.

The *architectural model of the building* and the *regime* turn out to be restrictive factors that are often correlated to the scores of the indicators. This proves that it is important to adjust the performance scores for restrictive factors.

Influenceable factors that correlate significantly to the indicators of the performance theme of safety are *management aspects* (such as the staff's workload, the extent of their autonomy, support by colleagues and styles of leadership), treatment styles used by staff, and the circumstances of detention of the prisoners (such as hygiene and contact with the outside world).

How can quantitative data be used to compare the performance of correctional institutions?

For this study, we have tested a model that calculates the expected performance score for every institution with the aid of restrictive factors, based on quantitative data. A comparison between the expected and the realised score makes clear how an institution performs and how the different institutions relate to one another. An order of ranking based on this model provides a considerably different picture than one that would have been based on the unadjusted figures. This proves that the model is useful.

In this way, it is possible to judge the performances of correctional institutions for every indicator while the local circumstances have been taken into account. Taken together, the indicators provide an all-encompassing picture of the performance of each institution with regard to each theme.

How can measuring errors be avoided?

While benchmarking, the risk of measuring errors and one-sided judgements can be considerably decreased by the use of different sources (such as registrations of the number of incidents, and surveys) and by measuring different aspects of one theme. With regard to the safety theme, for instance, one can measure how safe people feel *as well as* the number of registered incidents. This is what we have done in this study.

Another way to track down measuring errors is to check whether the measurements for different indicators are consistent with regard to content. This often turned out to be true.

Is benchmarking in the prison system effective?

By means of an ex ante implementation analysis, we have examined which consequences the introduction of benchmarking in the prison system may have. The methodology we have used consisted of three parts: an analysis of policy logic, a practice analysis, and a risk assessment.

The analysis of policy logic shows that it is to be expected, both in view of the experiences in other branches and countries and for theoretical reasons, that benchmarking may contribute to the intended goals.

The majority of prison governors in the expert panel reacted positively to the introduction of benchmarking, as a practice analysis has shown. The expectation is that improved comparisons of performances will generate 'a bright spot on the horizon', which will stimulate institutions to improve their performance and to learn from each other.

The prison governors in the expert panel do see risks, however. A rigid focus on particular performance indicators can make people focus their attention solely on what is measured, and can make the fear of making mistakes become preponderant. Yet, the governors think that such risks can be controlled by using different measuring methods and sources while benchmarking. Taking into account restrictive factors will be a help as well. Not all the institutions should be subjected to the same demands with regard to their performance; there should be some room for variation, enabling institutions to implement innovative measures. Above all, the prison management should always get the opportunity to throw more light on the figures and provide a context in discussions.

All in all, the ex ante implementation analysis provides a largely positive judgement of the introduction of benchmarking.

Conclusions

- In this study, we have found that it is possible to introduce benchmarking in the prison system and that this can contribute to the improvement of the performances of correctional institutions.
- Many of the technical and organisational preconditions for such an introduction are already being met. There is, for instance, an umbrella mission statement for correctional institutions, there are measurements that tell us something about the performances, and these measurements are carried out in the same way in all institutions. Between institutions, the scores for a number of indicators do actually turn out different, which is another precondition of benchmarking. It is possible to meet other preconditions without too much effort.
- Isolating the circumstances that cannot be controlled by management from those that it can influence makes it possible to make a better comparison between the performances of different institutions.
- It is to be expected that benchmarking will indeed result in improved performances, improve learning and clear accounts. Yet, from a scientific point of view, there is no guarantee that the instrument will be effective, nor is it certain that it will have the intended learning effect. There is a danger, for instance, that benchmarking will become no more than just a protocol, which will take up additional time from administrators and local managers without, in fact, yielding much. For this reason, it will be necessary to pay active and continuous attention to such side effects. In this context, a ‘learning attitude’ of those in charge should be the basic attitude.
- The regional prison governors should provide a sufficient amount of time and adequate means for innovations within the institutions. They should not immediately be too directing when performance comparisons become available. On the other hand, benchmarking should not be free of engagement; the sector management may stimulate and, if needed, guide the process of learning and improving. The willingness to change and learn from each other is decisive for the success of benchmarking.

Recommendations

The following recommendations serve to support a sensible introduction:

- Performance indicators should be connected with the organisation and the measurements should enable analysts to uncover differences between institutions. Against this background, it is recommended to examine the performance indicators periodically and to revise these when needed.
- More data should be gathered to provide the basis for an increase of the performance measurement regarding *human dignity* and *rehabilitation*.

- The possibility that *good practices* will not have the same effect in all institutions should be taken into account.
- Caution is called for when interpreting the order of ranking of performance scores: a lower order of ranking does not necessarily mean that an institution's performance is much worse in an absolute sense.
- It is necessary to keep investing in scientific research on factors connected to performance indicators that can be influenced.
- The standards determined for performance indicators should be dealt with consciously. An national average is not the ideal standard. There is a good chance that institutions that achieve above average will then slacken the reins, which will lower the average performance of all correctional institutions and will lead to a downward adaptation of the standard, which will generate a downward spiral.
- Benchmarking should be thoroughly embedded in the organisation of the prison system. See to the gathering of reliable data and let all persons in charge (in high and low positions) make active use of the instrument, in order to learn and to innovate.
- While forming their opinion, managers should always place figures regarding performance in a context and include 'the story behind the figures' in their considerations.