

Results

Workshop by the Federal Ministry of the Interior and the Forum of Federations on Benchmarking in Public Administrations in Federal States held in Berlin on 15 February 2008

On 15 February 2008 the Federal Ministry of the Interior and the Forum of Federations hosted an international workshop on benchmarking in public administrations in federal states. The workshop was based on proposals made by the German government in the framework of the second stage of the reform of Germany's federal structure to institutionalize benchmarking as a tool to compare performance in order to improve the quality of administrative action (effectiveness, efficiency, practicability, responsiveness to citizens' needs, staff satisfaction).

The following conclusions were drawn from the workshop:

1. Benchmarking in public administrations is a valuable tool to compare performance in a federal state. Administrations should be in a position to compare themselves with each another and to learn from best practices. Benchmarking requires permanent evaluation and adaptation to fulfil its functions. This also includes the definition of the relevant data basis.
2. Benchmarking studies can be conducted in public administrations and – possibly – in businesses performing the same/similar tasks. Not only the cooperation of experts from the administration, but also of external experts (methodological expertise) and – if possible – of partners from industry has proven useful.
3. The introduction of benchmarking requires the political will of the highest level of the federation and the constituent units. This political will was expressed by the decision taken by the Conference of Minister-Presidents on 15/16 December 2004 and by the statements of the Federal Government and the coalition parties on benchmarking as part of the second stage of the reform of Germany's federal structure.

4. If the constitution provides for close administrative and financial relations between the Federation, the constituent units and local authorities, all three levels must be included in the benchmarking procedure.
5. It is desirable to ensure that the benchmarking of public tasks not only focuses on costs, but also takes the effects of state action into account.
6. Organizational solutions for the federation and the constituent units are necessary to implement benchmarking.
7. Benchmarking results should be made public. A comprehensive discussion of benchmarking results and of possible conclusions conducted among the benchmarking partners should be an integral part of the benchmarking process.
8. The speakers reported that benchmarking so far has not been enshrined in their countries' constitutions. However, they felt that any support for the development of a sustainable benchmarking culture in federal systems was useful.

In detail:

After State Secretary Beus and the President of the Forum of Federations, Mr. George Anderson, opened the workshop, the following speakers presented their experience:

- Dr. Gottfried Konzendorf, desk officer, Project Group on the Second Stage of the Federalism Reform, Federal Ministry of the Interior, on the German discussion on benchmarking within the framework of the second stage of Germany's Federalism Reform,
- Jeff Tryens, consultant, former Executive Director of the Oregon Progress Board, on federal-state benchmarking in the United States with three examples from Oregon,
- John Wright, former Deputy Minister, Saskatchewan, Canada, on benchmarking healthcare in Canada,
- Dr. Reto Steiner, University of Berne, Competence Centre for Public Management, on benchmarking in the public administration in Switzerland, and
- Roger Wilkins, Head of Government & Public Sector Group, Citigroup (Australia and New Zealand), former Director-General of the Cabinet Office of New South Wales, Australia, on federalism and benchmarking in Australia.

The presentations specifically highlighted the following:

Dr. Gottfried Konzendorf (Germany, Federal Ministry of the Interior):

Benchmarking can be regarded as a method of learning

- to improve the quality of workflows, organizational structures and work results (input, output, outcome)
- based on a systematic comparison and exchange of experience of best practices
- between partners (organizations, local and regional authorities) performing the same or similar tasks or trying to solve similar problems.

Germany, like other countries, participates in international benchmarking studies, such as PISA or studies on better regulation. Benchmarking is also used for public tasks in Germany. In healthcare, for example, there are ten comparative studies entitled "Benchmarking healthcare". The tool is also used in other policy areas. The German consumer association "verbraucherschutz bundesverband" (vzbv) has already conducted a number of federal benchmarking studies on consumer protection policy. Other successful applications include the benchmarking studies of tax offices, local benchmarking studies of the IKO network and the benchmarking of city states.

When analyzed, these successful benchmarking processes produce the following results:

- Benchmarking in federal systems is possible and useful.
- Benchmarking helps improve the quality of administrative action.
- The methodological approach depends on the issue at stake. It is not possible to develop a uniform method that can be applied to all issues.
- Innovation partnerships between politicians, administrative staff and external consultants are recommended for the implementation of benchmarking to be successful.

Despite these positive examples there is still a huge demand for the use of benchmarking in the federal system of Germany. The decision taken by the Conference of Minister-Presidents on 15/16 December 2004 on comparisons of quality and performance has not been consistently implemented. Furthermore there is no sufficient data basis for comparative studies beyond the constituent states.

In order to extend benchmarking in our federal system a new section on “Cooperation” with the following article should be included in the Basic Law (the German constitution):

Art. Xx

(y) The Federation and the *Länder* shall conduct benchmarking studies to identify and promote the performance of their administrations and publish the results.

The Federation and the *Länder* should also take the decision to set up a benchmarking agency aimed at improving the implementation of benchmarking. This may be done by the Deutsches Forschungsinstitut für öffentliche Verwaltung Speyer (FÖV, German Research Institute for Public Administration), in which both the Federal Government and the Federal *Länder* cooperate and which is renowned for its methodological expertise. The following points must be taken into account, when assigning this responsibility to the FÖV:

- When conducting or assigning benchmarking studies, the FÖV will be supported by an advisory board of national and international administrative experts.
- The FÖV management board (members: Federation and the *Länder*) will decide on benchmarking topics upon the proposal of the members of Working Group VI of the Conference of Interior Ministers, the Federal Ministry of the Interior or other Ministers Conferences.
- The methodological basis and the results of benchmarking studies will be published by the FÖV in an appropriate form.
- The Federal Government and the Federal *Länder* are free to decide whether to participate in a benchmarking study and whether to draw any political conclusions.
- In December each year, a report will be published on the results of the benchmarking studies of the past year together with a plan for the coming year, which will be discussed by the Federal Chancellor and the Minister-Presidents of the *Länder*. This report will be forwarded to the Conference of Interior Ministers (Working Group VI) and possibly to other Ministers Conferences.

Jeff Tryens (USA, Oregon):

Most programs in the U.S. involve two, and often three, levels of government. Typically, the federal government provides significant funding and rules. State governments add additional funding and assume management responsibility for the program. Local governments, when involved, deliver services for the state government or oversee the work of those who do. State and local governments are fiercely protective of their policy-making prerogatives, balking at undue oversight from above. Benchmarking initiatives are further complicated by the sheer number of state and local governments involved. Over 80,000 local government units operate within the 50 states.

Perhaps because of this vast proliferation of sub-federal governments, federal initiatives that force performance-based competition between states are basically non-existent. The speaker was not aware of any federal program that pits states, head-to-head, in competition based on their respective measurable results.

This presentation will examine three types of results-focused federal programs from one state's perspective. The three types could be called: cooperative, soft prescriptive and hard prescriptive.

The cooperative example, known as the *Oregon Option*, was aimed at creating better results for Oregon by reworking intergovernmental relations. The soft prescriptive example, *Healthy People*, is based on a federal executive initiative designed to spur action at all levels of society by creating a well documented, target-based vision for the future. The hard prescriptive approach, *No Child Left Behind*, is based on a federal law requiring schools receiving certain types of federal funding to improve the academic performance of all their students.

This paper is based on the speaker's experience as former executive director of the Oregon Progress Board. As the group responsible for administering the state's results-based strategic vision, *Oregon Shines*, the Board has had some involvement in each of these approaches over its 19 year history.

PROGRAM DESCRIPTIONS

Cooperative Approach: Oregon Option

In 1994 a united front of Oregon state and local elected leaders convinced the Clinton Administration that Oregon was an ideal proving ground for trying out new approaches to government devolution based on shared objectives. Oregon was a

known policy innovator. It was far from the glare of East Coast media. And, perhaps most importantly, it had the Oregon Benchmarks: a statutorily-required set of measures focused broadly on the state's well being.

Over the next four years, Oregon and the federal government worked to improve results in three "clusters": child health, family stability, and workforce development. The initiative was considered a key cog in Vice President Al Gore's reinventing government machinery. Thousands of hours at all levels of government were invested in working out new ways to do business.

In preparation for this presentation, I polled ten key players involved in the *Oregon Option*. According to this admittedly small sample, their greatest hopes for the *Oregon Option* were: 1) achieve better outcomes in the clusters; 2) change the culture of federal-state or local relationships; 3) change the internal functioning of federal/state/local agencies. While the respondents felt that better outcomes had been achieved in the clusters, they felt that culture change and improved bureaucratic operation had been only somewhat or not at all successful. In short, the big picture results were disappointing.

Soft Prescriptive Approach: Healthy People

In existence for nearly three decades, this U.S. Department of Health and Human Services (DHHS) program is described as enabling "diverse groups to combine their efforts and work as a team." It provides data and national targets for 28 health topics with 467 specific objectives ranging from HIV to health education. Some prevention-related grants and no penalties are associated with the program. State coordinators are encouraged to share information with one another. The federal government also supplies standardized analysis and promotional materials. Comparative data on state progress toward the national targets is not a feature of the program.

Metrics included in program's first iteration, *Healthy People 2000*, were Oregon's measures of first choice when establishing the health-related Oregon Benchmarks. And the national targets set by DHHS guided the Progress Board's discussion regarding state targets. As federally-generated measures, the Healthy People metrics have had a certain cache with public health professionals. Oregon has participated in a national consortium that was formed to advance the goals of *Healthy People 2010*.

While still in existence, *Healthy People 2010* appears to have lost steam at the state level over the years. A former state "Healthy People" coordinator told me that

Healthy People 2010 drives no decisions at the state health department. Nevertheless, the federal government is now preparing *Healthy People 2020*.

Hard Prescriptive: No Child Left Behind (NCLB)

This federally mandated program, begun in 2002, requires all states to “have assessment systems, report disaggregated data and target federal resources to serve their neediest students” according to the federal secretary of education. Each state sets its own targets, without regard to national benchmarks, which schools receiving certain types of federal funding are required to meet for all students. Data is broken down into specific subsets of students for every participating school. Individual schools are subject to state and federal interventions, both positive and negative, as part of the initiative.

An interventionist, performance-based federal education policy is revolutionary in the U.S. where local elected school boards reign. Oregon state government, which has little control over what schools do, has welcomed the initiative. According to one state official, it provides “more and better opportunities to force change.” On the other hand, the federal government appears to treat NCLB as a one-size-fits-all tool with little collaboration among states or between states and the federal government. Again, state-to-state comparisons are not relied upon as part of act implementation.

While many problems remain, NCLB has established the federal government’s role in setting standards for educational performance. Perhaps its greatest achievement is that children are no longer passed from grade to grade without the requisite learning needed. The legislation is currently up for renewal in the U.S. Congress.

LESSONS LEARNED

Cooperative Model

1. *Desired outcomes must truly be shared.* Despite powerful federal political support for the *Oregon Option*, Oregon participants felt that the federal bureaucrats involved never really bought into the process. Sometimes mid-level federal bureaucrats acted as though the *Oregon Option* was a politically motivated free-pass that they had an obligation to thwart.
2. *Accountability for results must be carefully defined.* The Oregon Benchmarks, while important, were generally too high level to ascribe cause and effect relationships between government intervention and changes in benchmarks. A much more sophisticated set of metrics based on strategy implementation is needed.

3. *True systems change can only happen over the long haul.* Somehow, cooperative commitments must transcend terms of office. The *Oregon Option* was the product of a connection made between a state governor and the vice president of the United States. The departure of that governor in 1995 marked the beginning of the decline of the *Oregon Option* long before the end of the Clinton Administration, which sounded its death knell.

Soft Prescriptive

1. *Data alone doesn't do it.* Comparative data alone will generally motivate high performers to even higher levels but do little for others. Even in a field like public health, where most players are comfortable with data, the mere existence of regular, reliable, comparable data has done little to change the direction of health trends in Oregon in my opinion.
2. *Focus is critical.* A data-rich environment, allowing comparisons across a broad range of related issues, is some people's idea of heaven. It's my idea of hell. Imagine for a moment a yearly matrix of *Healthy People* data with 267 objectives across 50 (or even Germany's 16) states. Layering and priority-setting must be employed when tackling big issues.
3. *Soft incentives, like recognition for performance, are worthwhile.* Oregon uses sub-state level data to motivate counties to focus on benchmark-related issues by recognizing high performing counties and high improvement counties through an awards program. Free analysis is also provided upon request. More than once I've visited with a county leader whose office displayed the Progress Board's certificate of achievement.

Hard Prescriptive

1. *Even mandates should be soft around the edges.* Benchmarking programs should be carefully constructed to allow states to measure progress in ways that work. For instance, Oregon is changing its measurement system to track progress over the course of a classroom year rather than comparing achievement of last year's fifth graders to this year's.
2. *Mandates alone won't create a performance culture.* A school principle in Oregon recently told a state official that if he just had two fewer of a certain kind of student he could get out from under the *No Child Left Behind* requirements. What kind of a learning atmosphere is that for a disadvantaged child?
3. *The theory of change must be apparent.* Mandates are hard to swallow. Without a clear understanding of what the performance model is, even nascent supporters will balk. Is the model based on front line empowerment or top down authority, for instance?

Dr. Reto Steiner (Switzerland, Canton Berne):

In Switzerland, the federal administration, the cantons and the local authorities use benchmarking in very different ways:

- *Federal level:* Switzerland increasingly participates in benchmarking projects conducted by international institutions, such as the OECD's PISA studies. These comparisons are intended to assess the effectiveness of individual policy areas. Furthermore, the Federal Government promotes benchmarking projects between the cantons, e.g. in education and healthcare. This is implicitly required by legal acts laying down the principle of economical and efficient performance (e.g. Art. 43a of the Federal Constitution), which must be evaluated on a regular basis (e.g. Art. 170 of the Federal Constitution and Art. 36 of the Act on the Organization of the Government and the Administration). The Federal Government often acts as a facilitator, e.g. by tasking the Federal Statistics Office with the development of indicators, with organizing basic data or even with the interpretation of data. Federal offices are still hardly ever compared with one another.
- *Cantonal level:* At cantonal level, benchmarking projects are mainly initiated by the Federation or by conferences of directors (i.e. cantonal ministers' conferences). They usually focus on comparisons of outputs and outcomes between the cantons, rather than on comparisons of entire organizational units within cantons. Furthermore, cantons promote benchmarking at local level, because they want to make sure that the tasks assigned to local authorities are fulfilled in line with the requirements established by the cantons and that financial transfers are used efficiently.
- *Local level:* At local level there is a large number of benchmarking projects. They are initiated by the local authorities, by cantons or by other organizations, such as associations or consultancies. Since local authorities are usually small organizational units, there is a greater readiness to conduct comprehensive comparisons of organizations. In addition to systematically conducted projects, local authorities regularly carry out informal comparisons, as has been shown by a national survey of local authorities. The effects of such informal benchmarking should not be underestimated.

The following factors influence the success of benchmarking projects:

Success factor 1: Commitment of political leaders

- There must be a clear commitment to benchmarking, i.e. the support of political leaders and the heads of administrations.
- Staff members must be involved in the process (“onlookers must become participants”). Comprehensive information and training are a possibility to reduce existing barriers (barriers based on a lack of knowledge, abilities, wishes and legal norms). Fears of staff members as to whether a benchmarking process will eventually make staff redundant or involves individual sanctions, must be taken seriously.

Success factor 2: Integration and form of controlling

- Benchmarking is part of a comprehensive controlling and quality control strategy. This tool particularly makes sense if it is integrated in a new management and control model of the institution concerned. Quality management (e.g. the Common Assessment Framework) should be part of the administrative culture.
- For a systematic benchmarking the administration should know its products and design accounting in such a way as to enable absorption costing.

Success factor 3: Partners and objects of comparison

- The institutions chosen for a comparison in the first benchmarking study should be partners with similar tasks, i.e. with only minor differences. In a first stage, this would increase acceptance and facilitate work.
- Benchmarking partners from the private sector are an added benefit.
- Benchmarking can be initiated by institutions themselves, by the superior state level or by independent institutions. These three possibilities have their specific advantages and disadvantages. Cooperation with professional methodological experts and/or the creation of centres of expertise can be useful to benefit from know-how synergies.

Success factor 4: Applied procedures and tools

- The cost-benefit ratio should always be reasonable. Bureaucratic and oversized benchmarking procedures should be avoided.
- State benchmarking should not exclusively focus on costs (input orientation), but instead centre on processes and output/outcome.

Success factor 5: Evaluation, communication and improvement

- The internal discussion of results is essential.
- Results should be communicated in public and used as a basis for real improvements. Communication at the start of the project and after the

evaluation stage creates a certain pressure to act. Again, it is necessary to have the support of superiors and political leaders. This makes change more probable. It is important to pursue benchmarking continuously.

John Wright (Canada, Saskatchewan):

The use of comparable health care indicators for clinical, policy, program and research purposes is reasonably extensive in Canada. However, the use of benchmarking, based on best practice or medical evidence, is relatively new. In Canada, provinces provide health care services to 97 percent of the population. The federal government contributes less than 25 percent of the cost of provincial services through annual cash transfer payments to the provinces. Provinces are protective of their constitutionally assigned jurisdictions and do not welcome federal intrusions into their areas of responsibility.

Health care data has been traditionally collected by the provinces, Statistics Canada and the Canadian Institute for Health Information (CIHI). CIHI is a federal-provincial agency that is jointly funded and administered with a mandate to be a "source of unbiased, credible and comparable health information."

At the beginning of this decade, a rising sense of urgency to improve the quality and timeliness of health care led the federal government to inject additional funding into the system and in exchange the provinces were to provide increased transparency through comparable indicator reporting.

In 2000, 2003 and 2004, the First Ministers agreed to a series of health care accords that, among other initiatives, greatly increased the number of comparable indicators that were to be reported annually to the Canadian public. In addition, the 2004 accord required the provinces to expand comparable indicator reporting to four surgical areas and to produce evidence based benchmarks for each of these areas by December 2005. There is no legal basis for the accords and the benchmarking initiative.

In the development of the comparable indicators and benchmarks, no rigorous methodology was employed. A functional or collaborative approach was used with the following elements:

- Organize: A steering group of Deputy Ministers of Health was formed along with a working group that included *Statistics Canada* and CIHI officials.
- Plan: Clear definitions for comparable indicators were established and the necessary data infrastructure was defined. Similarly, a process for determining benchmark definitions for the four surgical areas was created.
- Collect Data: Best practices for data collection infrastructure across the country were examined and shared and implementation began in most provinces. Numerous challenges were encountered. A federal funding agency for health research was contracted to determine the evidenced based surgical benchmarks.
- Report Progress: Reports on up to 70 comparable indicators by province were made public in 2002, 2004 and 2006. Additionally, the evidence based benchmarks (8 in total) were published in late 2005.
- Analyze/Refine: Provinces were also required to produce by December 2007 a public plan on how they were going achieve the evidence based surgical benchmarks. No province met the deadline.
- Adopt Best Practices: Provinces are currently implementing best surgical practice methodologies and programs. Progress to achieving the benchmarks is being made. Each province has implemented one wait time guarantee relative to one of the benchmarks.

There was some early resistance to the development and publishing of the comparable indicators and benchmarks. Some provinces did not want to be compared to others. Also the cost and difficulty of collecting consistent data was deemed a barrier by some provinces. This resistance was largely overcome due to peer pressure, the public commitments of First Ministers, public pressure and pressure from various health care provider organizations.

Experience has shown that better upfront planning is needed. Work will continue with CIHI to ensure consistency in data definitions and collection and the sharing of best practices.

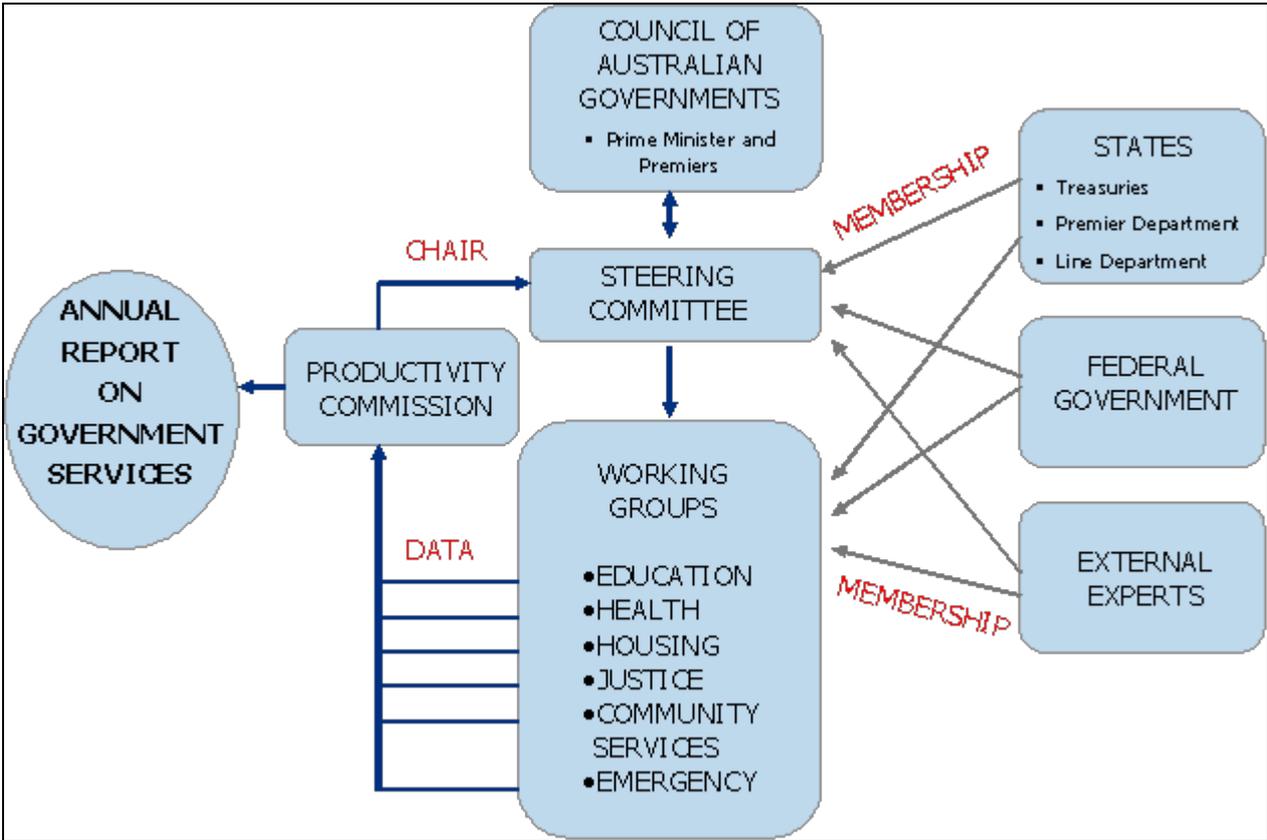
A benchmark research agenda is needed. A process to establish the next set of benchmarks including who decides, the order of priority and how the research will be undertaken needs to be established. Provinces need to look outside the health care system to establish certain benchmarks. The involvement of business schools, industrial engineering faculties and other partners is critical to providing robust benchmarks based on clinical evidence or best practice.

Public transparency and collaboration among the provinces have improved. The patient is on the road to recovery.

Roger Wilkins (Australia, New South Wales):

Since 1993, Australia has had a system of benchmarking government services delivered by State governments. There is no particular legal basis, only a decision of the Prime Minister and the Premiers.

There is an annual report on performance published by an independent Federal body called the Productivity Commission. The report covers services such as housing, health, education, policing, courts and justice, community services, and indigenous affairs. All of these services are delivered by the States. The Productivity Commission has been integrated into the following structure:



Progress has been made over the last 15 years in refining the system, but much still remains to be done in agreeing on and defining outcome measures.

State politicians sometimes find the report politically embarrassing, and the level of State support has been mixed. It has, however, enabled a degree of comparability and competitive federalism to drive efficiencies and innovation.

The new Australian Government (Labor elected in November 2007) has an agenda of wide ranging reform in government service delivery. It also works to get the federal system to work more efficiently. A focus on outcome measures and benchmarking is likely to see the reinvigoration/reform of the current benchmarking system.

A major feature of Australia's federal system is the level of vertical fiscal imbalance – the Federal Government raises most of the revenue; and there are large transfers to the States. This is an important reason why Federal governments have been able to encourage or coerce the States into benchmarking and reporting performance. But State Treasuries/Finance Departments and Premier and Treasurers have also found benchmarking a useful discipline in driving efficiencies.