Drivers for performance improvement originating from the Dutch drinking water benchmark

de Goede, Marieke; Enserink, Bert; Worm, G.I.M.; van der Hoek, Jan Peter

DOI
10.2166/wp.2016.125

Publication date
2016

Document Version
Accepted author manuscript

Published in
Water Policy

Citation (APA)
de Goede, M., Enserink, B., Worm, G. I. M., & van der Hoek, J. P. (2016). Drivers for performance improvement originating from the Dutch drinking water benchmark. *Water Policy*, *18*(5), 1247–1266. https://doi.org/10.2166/wp.2016.125

Drivers for performance improvement originating from the Dutch drinking water benchmark

Marieke de Goede[a]*, Bert Enserink[b], Ignaz Worm[c], Jan Peter van der Hoek[a]

a Delft University of Technology, Faculty of Civil Engineering and Geosciences, Section Sanitary Engineering, Stevinweg 1, 2628 CN Delft

b Delft University of Technology, Faculty of Technology, Policy and Management, Section Policy Analysis, Jaffalaan 5, 2628 BX Delft

c Ignaz Worm, Isle Utilities, Overboslaan 8, 2101 AM Heemstede

** Corresponding author. E-mail: j.a.c.degoede@tudelft.nl*

ABSTRACT

The Dutch drinking water sector has been benchmarked every three years since 1997, and the sector has significantly improved performance since then. Based on interviews with CEOs and financial managers of drinking water companies five drivers for improvement as a result of this benchmark are identified: 'learning effect', 'enhanced transparency', 'managed competition', 'avoidance of negative consequences' and 'personal honour of director'. Different developments have caused stagnation of further improvement: the variation on the benchmarked performance indicators has decreased, participation in the benchmark became mandatory for all Dutch drinking water supply organizations, it lacks a focus on the future, and participating organizations experience a high financial pressure. These developments decrease the influence of the drivers. Four possible new impulses for the benchmark are identified and their influence on the effect of the drivers is analysed. The two most promising new impulses are to make the benchmark adaptive and to involve consumers in the process of benchmarking, both have a positive influence on the effect of almost all drivers. This study contributes to the understanding of how benchmarking leads to improvement and to the analysis of the impact of design choices, leading to well-founded decisions for re-design of the Dutch drinking water benchmark.

Keywords | Adaptive benchmark; Benchmarking; Consumer involvement; Drinking water supply; The Netherlands

1. INTRODUCTION

Benchmarking originates from the private sector (Camp, 1989), it is a management instrument for comparison of performances of organizations. In the last two decades, the use of benchmarking in the public sector has increased (Blank *et al*., 2011), as a result of the fact that public organizations are often managed business-like manner, giving rise to the 'New Public Management' (Osborne, 2006). Benchmarking is seen as a powerful instrument for quality improvement in the public sector (Magd & Curry, 2003).

When benchmarking was adopted in the public sector, the process of benchmarking did not change, but the motivational forces differ from the private sector (Kouzmin *et al.*, 1999). There are different schools of thought about the use of benchmarking in the public sector. The first one sees benchmarking as a tool to stimulate collaborative learning among public sector organizations, which 'promotes organizational learning and enhances the transparency of the organizations' (Bolli & Emtairah, 2001; De Witte & Marques, 2007a). The second one presents benchmarking as an instrument for performance assessment, to help government control public services and create managed competition (van Helden & Tillema, 2005; Wynn-Williams, 2005).

Cabrera *et al*. (2011) have defined a benchmarking framework. They suggest use of 'comparative performance assessment' and 'performance improvement' as consecutive components of benchmarking. Performance assessment is the assessment of 'how efficiently or to which standard a certain utility operates'. Performance improvement is the phase in which the actual improvement takes place. It requires 'the participation of several utilities or benchmarking partners in order to gather additional information that will lead to the identification and adaptation of best practices'.

Many have studied the impact of benchmarks, however an important gap in available literature is the limited attention to the change processes through which benchmark findings or the comparative performance assessments are translated into improvement (Henry & Mark, 2003).

The results of the comparative performance assessment phase of a benchmark form incentives for change of behaviour in organizations. These incentives, in this article called 'drivers', stimulate organizations to improve (see Figure 1). A driver is defined here as an incentive that motivates people to improve their organization based on the performance assessment phase of benchmarking. This article focuses on the transition from the comparative performance assessment phase of benchmarking to the performance improvement phase in the Dutch drinking water benchmark. It answers the question through which drivers does the assessment of performance lead to improvement.

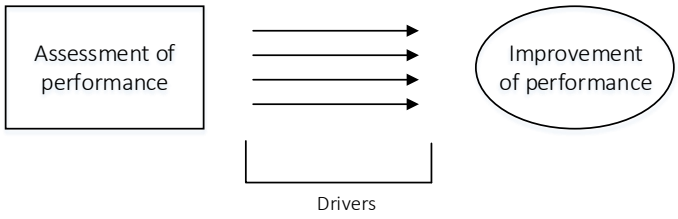

Figure 1: Drivers for performance improvement

This article studies the drivers of the Dutch drinking water benchmark. When these drivers are known it is possible to understand the origins of change of behaviour as a result of a benchmark study. This makes it possible to explain improvement and stagnation of improvement. The drivers have an additional function: they can be used as a framework to analyse the impact of design choices of the benchmark in a structured way, as shown in Figure 2. The results of a benchmark can lead through the identified drivers to change in behaviour of organizations. Stagnating improvement is a sign of loss of power of the drivers for change. It can be a reason for re-designing the benchmark. When the drivers of a benchmark are known, the influence of the new design choices can be analysed in a structured way by predicting their effect on the drivers.

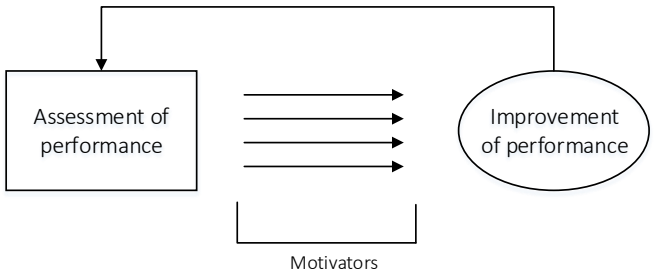

Figure 2: Drivers for performance improvement with feedback loop

An analysis of the drivers for improvement of the Dutch drinking water organizations as a result of benchmarking and their use to clarify trends based on the point of view of employees of Dutch drinking water utilities is not yet found in literature. Current literature about the Dutch drinking water benchmark focuses on the results of benchmarking (van Damme & Mulder, 2006; Braadbaart, 2007; de Witte & Marques, 2007b; Dijkgraaf *et al*., 2007; Schmitz & Dane, 2008; Dumaij & van Heezik, 2012). In literature an approach with similarities to ours is found (Henry & Mark, 2003). However, they have looked at different levels of change of behaviour, where we look at the drivers that lead to this change in behaviour.

This article focuses on the point of view of the Dutch drinking water utilities. This is a point of view not widely addressed in recent literature. Only one publication from the point of view of the water utilities is found in literature, the article of Neunteufel *et al*. (2010). However their focus is on the actual benefits of benchmarking in the Austrian water supply sector while our focus is on the drivers that lead to these benefits in the Dutch drinking water sector.

After a description of the history and present situation of benchmarking in the Dutch drinking water sector, the following questions are answered in this article:

1) Which drivers that stem from the Dutch drinking water benchmark have been responsible for performance improvement of the Dutch drinking water utilities?
2) Which drivers can be held responsible for stagnation of improvement and why?
3) How can improvement be encouraged again?

2. RESEARCH METHODS

Documents and interviews were used as data sources. Seven (out of the ten) Dutch drinking water organizations participated in semi-structured interviews. The interviewees were selected based on their knowledge about the use of benchmarking in their organization and on their availability and willingness to participate. From four organizations the directors were interviewed, one of them was assisted by his benchmark coordinator. From two organizations the benchmark coordinators were interviewed and from one organization the deputy director. In total, seven interviews were executed. The interviews took 1 to 1.5 hours.

It was decided to only interview employees of Dutch drinking water utilities, because this is the group of people that know if and how their organizations use the benchmark and if and how the results of the benchmark motivate for performance improvement, and in that way these interviewees contribute to the focus of this article.

As with all qualitative research, subjectivity might influence the collection of data. To improve the reliability of the research the interviews were recorded with the use of an audio-recorder and transcribed. The semi-structured interviews were guided by the following list of themes:

1. General questions about the use of the benchmark in their organization
2. Performance improvements as a result of the benchmark within their organization
3. Identification of drivers for performance improvement
4. Evaluation of the use of the results of the benchmark within organization
5. Possible new impulses for the benchmark

The transcribed interviews were first structured. The relevant answers were selected and subsequently labels were created. The labels were grouped and their impact was assessed. A hierarchy of the driver with the highest influence on performance improvement and the drivers with a lower influence was created.

This analysis of the interviews led to the identification of drivers that motivate people to improve their organization based on the outcomes of the performance assessment phase of benchmarking and the impact of these drivers. The identified drivers were verified by literature.

3. BENCHMARKING THE DUTCH DRINKING WATER SECTOR

3.1 The context of the Dutch drinking water benchmark

The Netherlands has ten drinking water supply organizations, with separated geographical supply areas. The drinking water organizations are Public Limited Companies with public parties as shareholders: municipalities and provinces (Huisman, 2002). The companies are managed by a board of directors and are supervised by a non-executive board, an independent accountant and the Inspectorate for Transport, Public works and Water Management (Steenhuisen, 2012).

Over the years, the number of water supply companies in the Netherlands has decreased. In 1985, the Netherlands counted about 90 water supply companies, nowadays 10 (Vereniging van Waterbedrijven in Nederland, Association of Dutch water companies, VEWIN, 2013). The scale increase was initially instigated by a change of the Drinking Water Act in 1975, to increase efficiency of the organizations and to improve the drinking water quality. This change of the Act gave the government more instruments to enforce drinking water supply companies to merge (de Witte & Marques, 2007b).

In the 1980s, a group of water supply companies, named COCLUWA, was involved in metrics benchmarking of its member companies (Schouten, 2009). In 1991 the first comparisons were made (Cabrera *et al.*, 2011), the results of this exercise were kept confidential (Blokland *et al.*, 2010). In the 1990s, a debate arose about privatizing the water supply companies in the Netherlands. The use of market forces in the water supply sector was expected to decrease the price of water by at least 10 per cent (Dijkgraaf *et al.*, 1997). The drinking water supply sector strongly opposed this privatization. The government desired more transparency of the costs of the activities in the drinking water sector. The drinking water supply companies tried to avoid direct government interference, and as a sign of goodwill started the voluntary drinking water benchmark in 1997 (De Witte, 2008). This benchmark concentrates on four themes: 'water quality', 'customer service', 'environmental impact' and 'financial performance' (VEWIN,

2013). Since 2012 participation in the benchmark has been mandatory for all Dutch drinking water organizations (Inspectie Leefomgeving en Transport, Human Environment and Transport Inspectorate, ILT, 2012).

3.2 Improvements as a result of the Dutch drinking water benchmark

Since the start of the voluntary benchmark in 1997 the performance of the Dutch drinking water sector has improved (van Damme & Mulder 2006; Braadbaart, 2007; de Witte & Marques, 2007; Dijkgraaf *et al.*, 2007; Schmitz & Dane, 2008; Dumaij & van Heezik, 2012).

Between 1997 and 2012 the real costs (taxes included) have decreased by 35% (without taxes 30%). The efficiency improvement per company ranges from 22% to 41% (VEWIN, 2013). Differences between companies are partly connected to differences in factors which are only minimally controllable, such as the use of surface water compared to the use of groundwater (VEWIN, 2013).

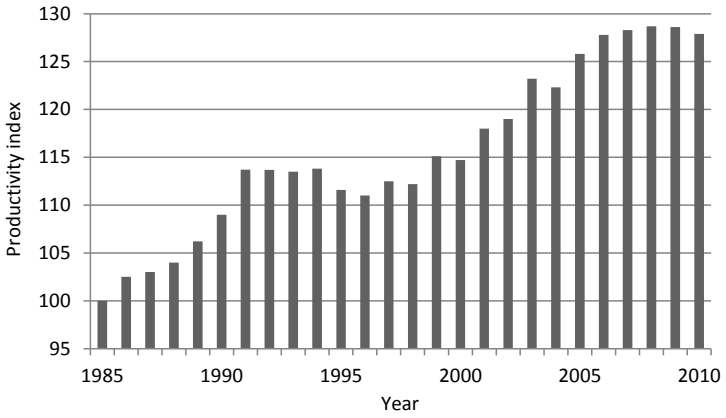

Figure 3: Productivity index (Dumaij, 2012)

Many sources suggest that benchmarking in the Dutch drinking water sector did contribute to its performance improvement, however it is impossible to indicate to what extent the benchmark has contributed to improvements. Dumaij and Van Heezik (2012) have researched the development of the productivity index over the years (Figure 3) and found a deviation from the trend of the productivity index starting in 2002. The benchmark was first executed in 1997, but only published in 1999. It seems apparent that the Dutch drinking benchmark did contribute to this increase of the productivity index and caused the deviation of the trend. However, the introduction of the Dutch drinking water benchmark cannot be seen separately from social developments, of which the political debate on liberalization of the Dutch drinking water sector is the most important one. The fear of privatization forced the Dutch drinking water utilities to increase their efficiency. It has stimulated the sector to reconsider their management (Dumaij & Van Heezik, 2012), the benchmark was an important tool to reach this. Drinking water companies themselves believe that the benchmark contributed to a large extent to the improvements of the sector. Table 1 shows the developments of the real costs (without taxes) per connection. The relative decrease of the real costs is the lowest after the first benchmark of 1997. This is caused by the fact that the results of the benchmark over the year 1997 were not published before 1999 (VEWIN, 2000). There was a limited time frame to take initiatives for improvements based on these results. The improvements between 1997 and 2005 are mainly a result of decreased operational costs, which can be influenced more easily than capital costs (Dijkgraaf *et al.*, 2005). As explained, Figure 3 shows the development of the productivity index from 1992

to 2010 (Dumaij & van Heezik, 2012). It shows an increase of the productivity from 2000 until 2008. After 2008 the productivity decreases.

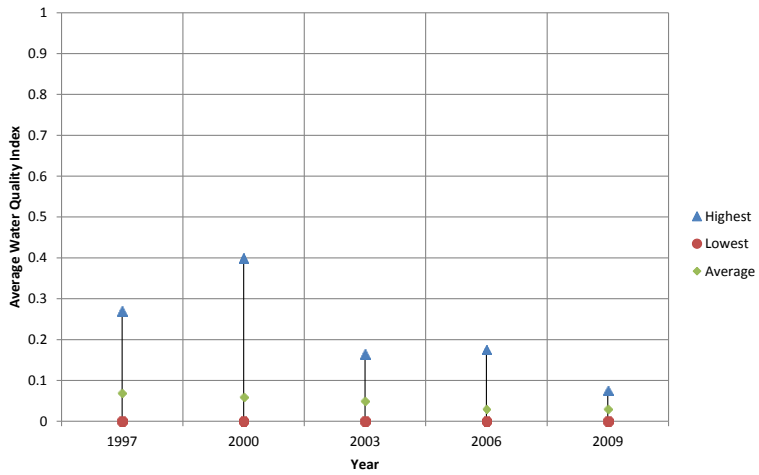

Figure 4: Water quality index* (Vewin, 2010)

While in 1997 the water quality index* of the different companies ranged from 0.00 to 0.28, in 2009 it ranged from 0.00 to 0.08 on a scale from 0 to 1, in which 1.0 indicates water that just meets the criteria from the Drinking Water Act, the closer to 0, the better the quality, see Figure 4.

The developments of the average drinking water tariff and the variation between the organizations is shown in Figure 5. A trend can be observed that the variation decreases over time until 2009, but increases again in 2012. The increase of 2012 is mainly caused by the abolishment of groundwater taxes (VEWIN, 2013).

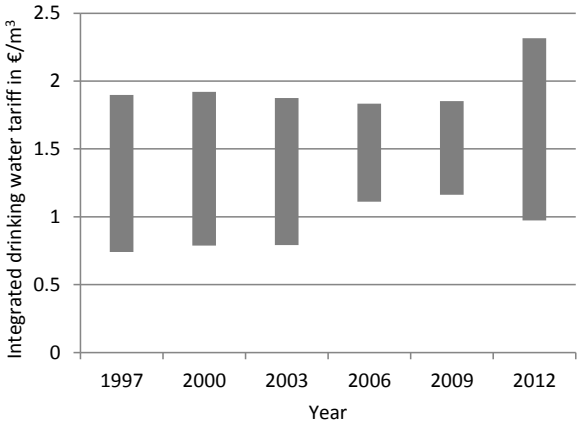

Figure 5: Drinking water tariffs (Vewin, 2013)

* Water quality index (WQI): the quality of drinking water expressed as WQI per group of parameters. The basis of the WQI calculation is the average relation between the measured value and the corresponding standard from the Waterleidingbesluit; (Vewin, 2010)).

Another aspect evaluated in the benchmark is the grade that consumers give the Dutch drinking water supply organizations for their water quality over the years. The average perception of the performance of the sector has increased over the years, and the variation between the lowest and highest scoring organization has decreased. In 2003 the grade ranged between 7.3 and 8.3 and in 2012 between 8.1 and 8.5 (on a scale of 1 to 10, where 10 is the highest possible score). The grade for customer service follows the same development, in 1997 the grade varied between 7.3 and 7.9 and in 2012 between 7.6 and 7.8 (VEWIN, 2013). The variation thus decreased from 0.6 points to 0.2 points.

Table 1: Development of the real costs in Euro per connection

	Real price per connection	Decrease
1997	€ 239	
2000	€ 235	1.7%
2003	€ 213	9.4%
2006	€ 198	7.0%
2009	€ 179	9.6%
2012	€ 168	6.1%

4. DRIVERS FOR PERFORMANCE IMPROVEMENT IN THE DUTCH DRINKING WATER BENCHMARK PERIOD 1997-2012

The previous section showed that the Dutch drinking water sector has improved since the start of the benchmark, but improvements are getting smaller.

In this section the drivers that motivate for performance improvement are identified and their degree of influence on the described improvements is estimated.

Based on the interviews, five drivers for improvement for the Dutch drinking water benchmark are identified:

1. Learning effect
2. Enhanced transparency
3. Managed competition
4. Avoidance of negative consequences
5. Personal honour of director

Six of the seven interviewees indicated that learning from other organizations has played a role in improvement. This driver is also named frequently in the literature (for example Bolli & Emtairah, 2001; van Helden & Tillema, 2005; de Witte & Marques, 2007b). They believe that this learning effect was mainly present during the first few benchmarks. During these first benchmarks 'best practice' sessions were organized, in which information was exchanged between organizations.

The desire to increase transparency can form a reason to start benchmarking (Bertzbach *et al*., 2012), this was also one of the main reasons the Dutch drinking water benchmark was started (De Witte, 2008). From the institutional point of view, organizations have a need to preserve support from stakeholders (van Helden &

Tillema, 2005). Benchmarking creates transparency and from this transparency follows accountability to other stakeholders. All interviewees believe that improvement driven by the enhancement of transparency and accountability plays a very important role. This is affirmed by Witte (2008): 'Even though the drinking water supply organizations are regional monopolies, it seems that the public opinion has a rather large influence on their behavior'.

All interviewees experience the third driver, managed competition. 'Managed competition' finds its origin in the economic theory of public sector benchmarking (van Helden & Tillema, 2005). Drinking water supply organizations are natural monopolies. In markets with competition, markets give organizations information about the demand for products and services, the objectives and their performance. If a market consists of natural monopolies, there is no consumer choice. This implies that negative performance gaps are not a direct threat to the survival of public organizations (Nutt & Backoff, 1993; Dijkgraaf *et al.*, 1997). Benchmarking is a form of managed competition, it offers a substitute for consumer choice, and it enables consumers to compare relevant performance characteristics of different suppliers (van Helden & Tillema, 2005).

The interviewees believe that every organization is different, and comparison is not always fair. Most organizations choose one or more spearheads, on which they desire to reach a certain position in the benchmark, for example be one of the three best performing organizations. These organizations are driven by managed competition on those specific spearheads.

The fourth identified driver is the 'avoidance of negative consequences'. The risk with a monopoly is that the supplier misuses his exclusive suppliers' rights, which may harm society. To guarantee economic efficiency, government control is needed (Png, 2002). Benchmarking can be used for government control. Governments can impose fines or other consequences, such as determination of pricing, based on the outcomes of benchmarking. The incentive for improvement that follows from such measures, 'avoidance of negative consequences', is our fourth driver. The Dutch drinking water benchmark system does not include direct consequences based on the results of the benchmark, such as imposing fines for organizations with a low performance. However all interviewees indicate fear of further involvement of the government in the operation and management of their organizations. The fear of further government involvement, with the possibility of creating negative consequences for low scoring organizations, does function as a driver for improvement.

The fifth driver for improvement is the fact that directors of the organizations participating in the benchmark, do not want to be last, as a matter of personal honour. This fifth driver is called 'personal honour of director'. According to Adam Smith (The Wealth of Nations, 1976), each individual strives to become wealthy or, in other words, to pursue his own self-interest. One of the interviewees explained it as follows: 'Managers always experience a certain tension, in which their own position and keeping that position plays an important role'.

All interviewees confirmed that this driver has an influence. However, it is difficult to indicate its degree of influence. For the interviewees who are not directors, it is difficult to estimate the influence of this driver, simply because they do not have the knowledge. For the directors themselves we have to take into account the 'self enhancement' bias, the tendency of people to evaluate themselves more positively than others (de Cremer, 2012). From a moral point of view, a director's own interests should be secondary to the interests of the organization.

The interviewed directors indicated that the personal honour of the director plays a role, but they believe that this role is small for their own organization.

Based on the interviews the degree of influence of the drivers from 1997 until 2012 is determined and shown in Table 2.

Table 2: Drivers for improvement and their influence: 0 no influence, + moderate influence, ++ reasonable influence, +++ high influence

Driver	Influence on improvement
1. Learning effect	++
2. Enhanced transparency	+++
3. Managed competition	+
4. Avoidance of negative consequences	+
5. Personal honor of director	+

5. STAGNATION OF FURTHER IMPROVEMENT

In this section reasons for stagnation of further improvement are identified, based on the interviews. The interviewees indicated that they experience several issues with the benchmark, that may cause stagnation of further improvement. These issues can be grouped into four categories:

- Drinking water organizations believe that the variation between the participating organizations has decreased to a level where the benchmark is losing impact
- Participation in the Dutch drinking water benchmark has become mandatory for all Dutch drinking water supply organizations which decreases voluntary learning
- The Dutch drinking water benchmark has no focus on the future
- The current system pays too much attention to financial aspects

These issues and their influence on the identified drivers is explored in this section. This is not a complete list of reasons for stagnation, this list gives an overview of the issues that were mentioned in the interviews. Other explanations for stagnation are possible, for example the fact that the Dutch drinking water utilities have now picked all the low hanging fruits, and further improvement becomes difficult. However, here we only focus on the issues from the point of view of the drinking water utilities as identified during the interviews.

5.1 Less variation

The decrease of the variation in the performance indicators in the Dutch drinking water benchmark has been caused by two factors. First, the number of companies decreased. In 1997 23 drinking water supply organizations were active in the Netherlands (Braadbaart *et al.*, 1999), nowadays there are only 10 (VEWIN, 2013). Secondly, the variation decreased as a result of the benchmarks. It has been proven that organizations take initiative to improve as a result of a benchmark, when the results of the benchmark disappoint (van Helden & Brouwer, 2005). Benchmarking therefore leads to convergence to an average level of performance, resulting in decreased performance variation (de Bruijn, 2002; van Helden & Tillema, 2005).

For all performance indicators discussed in the previous section, the variation has decreased, except for the tariffs. An aspect that may be even more important than the actual remaining variation, is the fact that the drinking water

supply companies are convinced that the remaining variation is mainly (if not completely) caused by external factors, that cannot be influenced by the organization itself (Source: Interviews). Examples are the difference between companies that use groundwater and companies that use surface water, the density of the supply area, the size of the company and the subsidence of the surface in the supply area. Regardless of sharing a common business and perhaps many other circumstances, participating companies are often very different (Cabrera, 2008). As the number of participating companies decreases, 'it becomes easier for the remaining utilities to invoke exogenous influences when they perform less efficiently according to the benchmark study' (van Helden & Tillema, 2005).

The remaining variability is only used to a little extent to learn. Two out of the seven interviewed organizations indicated that they still learn from other organizations. The other five believe that the learning effect has decreased to a minimum or that the learning effect has never existed.

Performance measurement requires not only comparability but also variability, in order to discriminate good from bad performance (Meyer & Gupta, 1994). If the variability decreases, the discriminative power between good and bad decreases. As explained, benchmarking itself decreases performance variation, because it often leads to convergence to an average level of performance. The impact of performing less efficiently decreases. This means that drivers that are based on a clear distinction between good and bad exert less power ('managed competition', 'avoidance of negative consequences' and 'personal honour'). In this study no influence of the development 'less variation' on the driver 'enhanced transparency' is found.

5.2 Mandatory character of benchmark

From 2012, participation in the drinking water benchmark has been mandatory for all drinking water supply companies. Table 3 shows the percentages of the Dutch drinking water utilities that participated in the Dutch drinking water benchmark. From 2006 all utilities participated, while the benchmark became mandatory by law only in 2010. This indicates that the mandatory character did not influence participation of the utilities in benchmarking. The drinking water sector itself agreed to self-imposed mandatory participation in 2004, prior to the Drinking water law that became effective in 2009 (Dumaij & van Heezik, 2012).

Table 3: Percentage of drinking water utilities participated in benchmark

year	Percentage of drinking water utilities participated in benchmark*
1997	85% (VEWIN, 1999)
2000	90% (VEWIN, 2001)
2003	81% (VEWIN, 2004)
2006	100% (VEWIN, 2007)
2009	100% (VEWIN, 2010)
2012	100% (VEWIN, 2013)

*percentage is given in representing number of connections in the Netherlands

All interviewed companies consider the fact that participation is now enforced by law a negative development (Source: Interviews). They believe that something that was originally an improvement driven initiative became an instrument for regulation. The tension that was created by the benchmark, and that was an incentive for companies to improve, decreased (Source: Interviews).

Using learning and regulation objectives at the same time, can lead to too much attention on one of the objectives (de Bruijn, 2002). Organizations can become so focused on illustrating their performance level is reasonable, that the learning objective of benchmarking is put under pressure (van Helden & Brouwer, 2005). This is what happened according to the interviewees in the Dutch drinking water benchmark. The mandatory character of the benchmark has decreased the learning effect of it. The new policy of the Ministry of Infrastructure and the Environment has attempted to deal with this aspect by introducing a new incentive for improvement: all companies have to write a plan for improvement (van Geel, 2006). The companies themselves are quite sceptical about these plans which they have to submit 6 months after the benchmark is produced (Source: Interviews).

When benchmarking has a mandatory character, the chance that undesired consequences occur increases. This is called the paradox of the increased perverse effects (de Bruijn, 2002). Professionals strengthen their attempts to show that their performance is good. Drinking water companies know the aspects on which they are judged, and do their best to perform well, possibly specifically on these judgmental aspects. The measures (in the sense of indicators) become targets and as Goodhart's law stated (Goodhart, 1975): 'When a measure becomes a target, it ceases to be a good measure'.

When benchmarking is mandatory, it can lead to more strategic behaviour of the participating organizations. Strategic behaviour has a negative influence on the driver 'transparency', because it is focused on showing good performance instead of executing good performance. According to the drinking water organizations this is not happening in the Dutch drinking water sector (Source: Interviews). However, the mandatory benchmark has only been executed once, so it might change in the future, especially when it turns out that the results of the benchmark will have consequences.

Mandatory benchmarking may lead to a situation where organizations become less willing to share knowledge with other organizations, because they compete with each other based on performance (de Bruijn, 2002). To date the Dutch drinking water supply companies are still open to each other. There is still the possibility to learn from the best in class. However, when a company asks another company for specific data for example about cost prices of certain resources this is not always shared. The extent of sharing knowledge may become lower in the case of a mandatory benchmark (Source: Interviews). This will have a negative influence on the driver 'learning effect'.

With the mandatory character of the benchmark, the degree of competition may increase, with the consequence that the influence of the drivers 'managed competition' and 'personal honour' increases. The mandatory benchmark intensifies government involvement in the sector, this increases the fear of even more government involvement with an increase of the driver 'avoidance of negative consequences' as a result.

5.3 No focus on future

The average time perspective of the 'foreseeable future' of Dutch water professionals is 2.8 years (Segrave, 2014). Compared to other countries this is relatively short. The short time perspective of Dutch water professionals is, according to Segrave (2014), a result 'of the sense that the fundamental purpose or mission has been achieved and the focus is on pragmatic management maintenance and gradual optimization'.

This short term view is also visible in the drinking water benchmark. The Dutch drinking water supply companies have expressed that they believe the benchmark only has attention on the short term, and neglects the long term.

The short time perspective of water professionals and the drinking water benchmark does not match with the nature of drinking water production. Drinking water production takes place in a complex environment and investments in the sector often have a depreciation period of 30 to 50 years.

Examples of how focus on the future can be incorporated in a benchmark include performance indicators for risk analysis, asset management, sustainability and innovation. These themes contribute to the focus on the future of drinking water utilities. Because the benchmark is seen as an instrument without a focus on the future, organizations are not able to compare and learn from each other on the aspects of future focus based on the benchmark. This means that the learning effect of the benchmark is lower than it may be. Future information also plays an important role for the driver 'transparency'. Currently the 'future' of drinking water supply organizations is not part of the benchmark, this means automatically that the benchmark does not create transparency and accountability on these future aspects.

The fact that the benchmark has no focus on the future, seems to have no proven influence on the drivers 'managed competition', 'avoidance of negative consequences' and 'personal honour of director'.

5.4 Financial pressure

The drinking water supply companies believe that the sector pays too much attention to financial aspects (Source: Interviews). Companies believe that the price of drinking water in the Netherlands is already very low. An average Dutch household pays around 200 euros for treated drinking water per year, this excludes a charge for waste water discharge. This is on average 0.6% of the total household budget (VEWIN, 2013). Even though some drinking water organizations value drinking water quality and continuity of supply higher than the financial aspects, the current regulatory system and institutional interactions force them to value the financial aspect as very severe (Source: Interviews).

Benchmarking, especially when the financial pressure is high, can obstruct innovation (de Bruijn, 2002). Innovation means exploration of the unknown and thus increases the risk that the results are lower than expected. Benchmarking rewards reproduction of the known. The financial pressure thus reduces the learning effect.

If the feeling of financial pressure increases, the degree of competition between the drinking water supply organizations may increase with the consequence that the influence of the drivers 'managed competition' and 'personal honour of director' increases. Increased financial pressure may also lead to more strategic behaviour, with a negative influence on the driver 'transparency'. In this research no influence of financial pressure on the driver 'avoidance of negative consequences' is found. This does not mean that there is no relation between the two, one could imagine that increased financial pressure may increase the avoidance of negative consequences, because this again may have financial consequences. However, this relation does not follow from this research.

5.5 Overview of influence of developments on the drivers

Table 4 gives an overview of the influence of the above mentioned four issues on the five drivers.

The drivers that had the most important share in stimulating improvement from 1997 until 2012 'enhanced transparency' and 'learning effect' (Table 2), are the drivers that are most negatively influenced by all four developments (Table 4). The influence of these drivers has decreased, on the other hand, some drivers have been

stimulated by the developments. The driver 'managed competition' is stimulated by two of the four developments, as well as the driver 'personal honour of director'.

Table 4: Overview of consequences of the issues 'less variation', 'mandatory character', 'no focus on future' and 'financial pressure' on the five drivers: ++ very positive; + positive; 0 neutral; - negative; -- very negative

Developments / Driver	Less variation	Mandatory character	No focus on future	Financial pressure
1. Learning effect	--	--	-	--
2. Enhanced transparency	0	--	-	-
3. Managed competition	-	++	0	+
4. Avoidance of negative consequences	-	+	0	0
5. Personal honor of director	-	+	0	+

The most important negative influence follows from the development 'less variation'. This development has a negative influence on four out of five drivers. The mandatory character has an adversative influence on the different drivers. It negatively influences 'learning effect' and 'enhanced transparency', but it positively influences 'managed competition', 'avoidance of negative consequences' and 'personal honour of director'. From 1997 until 2012 the degree of influence of these three drivers was relatively low (Table 2).

6. POSSIBLE NEW IMPULSES AND THEIR CONSEQUENCES

In the previous section the drivers that have lost power have been identified. The drivers have been used to explain improvement and stagnation of further improvement in the Dutch drinking water sector. The drivers have another function: they can be used as a framework to analyse the impact of design choices of the benchmark in a structured way, as shown in Figure 2. Stagnating improvement is a sign of loss of power of the drivers for improvement. It can be a reason for re-designing the benchmark. When the drivers of a benchmark are known, the influence of the new design choices can be analysed in a structured way by predicting their effect on the drivers. This section uses the drivers to investigate possible new impulses for the Dutch drinking water benchmark. Four new impulses have been investigated. The impact of these new impulses on the five drivers is assessed based on extrapolation of the findings from the interviews and is shown in Table 5. These impulses are selected based on the interviews.

Table 5: Overview of the new impulses on the five drivers: ++ very positive; + positive; 0 neutral; - negative; -- very negative

New impulses / Driver	Adding fixed new themes	Adaptive benchmark with changing new themes	Increase number of participants	Consumer involvement
1. Learning effect	++	+++	+	+
2. Enhanced transparency	++	+++	+	+++
3. Managed competition	+	+	--	+
4. Avoidance of negative consequences	0	0	0	+
5. Personal honor of director	+	+	0	+

6.1 New themes

6.1.1 Fixed new themes

The first identified new impulse for the Dutch drinking water benchmark is the addition of new themes to benchmark that introduce variability again. As a consequence the drivers that were negatively influenced by the development 'less variation' (Table 4), are expected to be stimulated: 'learning effect', 'managed competition', 'avoidance of negative consequences' and 'personal honour of director'. With new themes organizations give insights in new aspects of their company, thus the transparency will be increased.

The success of this new impulse depends on the selection of new performance indicators. The selection of themes for a benchmark is a very challenging task (Van Peursem *et al.*, 1995), and is described as one of the most crucial aspects of benchmarking (Trosa & Williams, 1996). The new performance indicators should be experienced as important themes by the drinking water organizations. Based on the interviews they should focus on the future orientation of the organizations. All interviewees believe that the quality of the benchmark will improve if future orientation is added to the benchmark (Source: Interviews).

It is also an idea to create a benchmark with fewer performance indicators for the Dutch drinking water benchmark. However, the currently used indicators are an essential part of the operational management of water utilities. Therefore, comparisons of these indicators remain interesting. It is a possibility to let the current performance indicators stay part of the mandatory benchmark, and to introduce an additional voluntary benchmark with new indicators.

6.1.2 Adaptive benchmark with changing themes

An adaptive benchmark is a benchmark in which the themes are not fixed, but may change over time, based on developments that influence the Dutch drinking water sector. New themes are to be added, and 'exhausted' themes are to be deleted.

An adaptive benchmark is expected to have, in the first instance, the same influence as adding new themes. However, with an adaptive benchmark this influence will keep existing, because it is an evolving process. When adding a new performance indicator, organizations may improve on this indicator. However the lowest scoring organizations improve more than the best scoring organizations which leads to convergence to an average level of performance again, with the consequence of stagnation. An adaptive benchmark pre-empts this with the ongoing possibility of deleting exhausted themes and adding new themes.

Making a benchmark adaptive is expected to have a positive influence on four of the five drivers. The new themes have variability again, which improves the driver 'learning effect'. New themes will create transparency on aspects of organizations that were previously not present or out in the open. The themes on which competition occurs are broadened, which is expected to influence the drivers 'managed competition' and 'personal honour of director' positively. An adaptive benchmark is expected to have no influence on the driver 'avoidance of negative consequences'.

When desiring an adaptive benchmark, adding and deleting themes will be an ongoing process. Themes that are benchmarked should be allowed to be deleted, because the circumstances change over time and themes might

lose impact and become exhausted. It would contribute greatly to the use and development of the benchmark, if the choice for introducing new themes and deleting old ones can be made in a transparent and objective way.
To be able to add new themes, a method needs to be developed to analyse future trends in a structured way, leading to transparent choices for adding new themes to the benchmark. To be able to delete exhausted themes, specific rules need to be generated on when a certain theme is exhausted and should be deleted from a benchmark. A theme may be exhausted for different reasons, e.g. if the learning effect is over or if it is not interesting for transparency or managed competition purposes. This may happen for instance if all of the remaining variation is caused by external factors on which the participants have no influence.

6.2 Increase number of participants

Another possible new impulse for the benchmark is to increase the number of participants in the benchmark, by cross-sector benchmarking or by international benchmarking. The interviewees have different opinions on increasing the number of participants. Some find it interesting, while others do not. The most important argument against it, is the idea that the comparability of the participants decreases with international benchmarking as well as with cross-sector benchmarking.
Kouzmin *et al*. (1999) believe that public sector benchmarking should only be used between public organizations that have very similar goals and other organizational characteristics, 'so that actors actually perceive differences or qualitative improvements in delivering similar services to constituencies' (Dixon & Kouzmin, 1994). If the comparability decreases as a result of an increased number of participants, the influence of the driver 'managed competition' is expected to decrease.
Cross-sector benchmarking and international benchmarking can have an added value; however this added value is only reached if it is used for specific themes on which participants are comparable. For cross-sector benchmarking examples of such themes are climate footprint, employee satisfaction, customer satisfaction and complaints procedures.
With more participants for specific themes, the variability increases with the consequence that the effect of the driver 'learning effect' can increase again. The transparency may also increase, because the position of an organization gives more information when more organizations participate. For a benchmark over the company-wide performance, cross-sector and international benchmarking seems to be less suitable.
Increasing the number of participants is expected to have no influence on the drivers 'avoidance of negative consequences' and 'personal honour of director'.

6.3 Consumer involvement

Consumer involvement in the benchmark means that consumers have a voice in the selection of themes that will be benchmarked, and/or in the weight that is attributed to the different benchmarked themes. This can be implemented in different ways, for example by creating customer focus groups or surveys to collect information. As the definition of performance indicators is a difficult task, the corresponding performance indicators should be determined by specialists.

The current Dutch governance system has multiple means of consumer involvement. Six of the ten Dutch drinking water utilities have drinking water panels, in which a group of customers has regular meetings and can express their opinion. In the current benchmark consumers are consulted about their experience of the water quality and their experience of the customer service. However, they have no voice in determining which themes are most important for them and should be benchmarked. The drinking water market has a monopolistic character, which means that it is not possible for consumers to express a voice by choosing a product from a different company.

Drinking water supply is an important public service. Evaluation researchers have argued that the ultimate goal of evaluation of public services should be social betterment (Lawrenz *et al.*, 2007). Social betterment is 'a closer correspondence between perceptions of actual and desired standards of public services' (Boyne, 2003). A good benchmark should thus decrease the gap between the actual and desired standards of the service.

The current system assumes that shareholders of the Dutch drinking water supply organizations adequately manage public values. However Steenhuisen (2012) explains that drinking water tariffs and efficiency are the dominant aspects that public shareholders of drinking water organizations discuss during shareholder meetings. Research shows that Dutch consumers have little to no interest in the developments related to regulation of prices in the sector, and consumers show a relative large willingness to pay for societal projects like nature preservation and development aid (Pricewaterhouse Coopers, 2008). Differences can be expected between what consumers desire and how this is arranged in the institutional system.

Consumer involvement is expected to have a positive influence on all five drivers. It will create a situation in which the benchmark measures the aspects that must be measured according to consumers. Transparency is obtained on aspects on which consumers actually desire transparency. The learning effect will increase, because drinking water supply organizations will be better informed about the preferences of consumers and can improve on those aspects. Consumer involvement is also expected to have a positive influence on 'managed competition' and 'personal honour of director' because these then will take place on the 'right' aspects. The driver 'avoidance of negative consequences' may increase. Measures will be more effective, because they are based on themes that are most important for consumers.

6.4 Overview of influences of new impulses

Table 5 shows an overview of the perceived influences of the possible new impulses on the five drivers. The table shows that all impulses are expected to influence the drivers. The impulses 'make benchmark adaptive' and 'consumer involvement' are believed to have the most positive influence on the drivers. To make the benchmark adaptive a model has to be created to objectively decide on adding new themes and deleting exhausted themes. To realize consumer involvement a method should be developed on how this involvement can be best expressed.

7. CONCLUSIONS

This paper identified drivers for performance improvement originating from the Dutch drinking water benchmark from the point of view of employees of Dutch drinking water utilities. With these drivers, previous improvements in the Dutch drinking water sector are explained, and reasons for stagnation of further improvement are clarified. The identification of these drivers contributes to the understanding of how benchmarking leads to improvement

in the Dutch drinking water sector and it creates the opportunity to analyse the impact of design choices of this benchmark in a structured way, leading to well-founded decisions for re-design of the Dutch drinking water benchmark.

The two most promising new impulses for the Dutch drinking water benchmark following from this research are to make the benchmark adaptive and to involve consumers in the process of benchmarking. An adaptive benchmark requires a method that makes transparent and objective choices for performance indicators. Such a method is not yet available, and should be developed before an adaptive benchmark can be made. However, if it is possible to develop such a method, it is believed that an adaptive benchmark will have a positive influence on the effects of four of the five drivers.

The second suggested new impulse is the involvement of consumers in the Dutch drinking water benchmark, meaning that consumers will have a more direct voice in the selection of benchmarked themes or in the weight attributed to them. Consumer involvement is expected to have a positive effect on all five drivers. Further research is needed to clarify which form of consumer involvement will be best applicable.

This paper has focused on the Dutch drinking water benchmark. The identified drivers are specific for this sector and selected based on the point of view of employees of Dutch drinking water utilities. However, the paper also forms a starting point for a model of drivers for change of behaviour as a result of benchmarking in general. Additional research is needed to be able to generate such a model.

References

Bertzbach, F., Franz, T. & Möller, K. (2012). How to achieve and prove improvement – 15 years of experience in German wastewater benchmarking. *Water Science & Technology*, *65(4), 661-668*.

Blank, J., Felsö, F. & de Groot, H. (2011). Beter benchmarken - Van cijfers naar informatie. *Tijdschrift voor Public Governance, Audit & Control, 9(6),* 4-9.

Blokland, M., Schouten, M. & Schwartz, K. (2010). Rejuvenating a veteran benchmarking scheme: benchmarking in the Dutch drinking water sector. *Journal of Competition and Regulation in Network Industries, 11(2),* 130-152.

Bolli, A. & Emtairah, T. (2001). *Environmental Benchmarking for Local Authorities: From Concept to Practice*. European Environment Agency, Copenhagen.

Boyne, G. A. (2003). What is public service improvement? *Public Administration, 81(2)*, 211-227.

Braadbaart, O. (2007). Collaborative benchmarking, transparency and performance. *Benchmarking: An International Journal, 14(6)*, 677-692.

Braadbaart, O., Blokland, M. & Hoogwout, B. (1999). Evolving market surrogates in the Dutch water supply industry: investments, finance and industry performance comparisons. In: *Private Business, Public Owners. Government Shareholdings in Water Enterprises*, M. Blokland, O. Braadbaart & K. Schwartz (eds.), Minsterie van VROM, The Hague, pp. 81-92.

Cabrera, E. (2008). Benchmarking in the water industry: a mature practice? *Water Utility Management International*, *3(2)*, 5-7.

Cabrera, E., Dane P., Haskins, S. & Theuretzbacher-Fritz, H. (2011). *Benchmarking Water Services – Guiding Water Utilities to Excellence*. IWA Publishing, London.

Camp, R. C. (1989). *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*. ASQC Quality Press, Wisconsin.

de Bruijn, H. (2002). *Managing Performance in the Public Sector*. Routledge, London.

de Cremer, D. (2012). *When Good People do Bad Things*. Acco, Antwerpen.

De Witte, K. (2008). Regulering zonder regulator – Regulation without a Regulator. *Economische Statistische Berichten, 93(4547),* 692-693.

De Witte, K. & Marques, R. C. (2007a). *Designing incentives in local public utilities, an international comparison of the drinking water sector*. CES Discussion Paper Series.

de Witte, K. & Marques, R. C. (2007b). Regulering brengt efficiëntie in de drinkwatersector – Regulation in the drinking water sector leads to efficiency. *Economische Statistische Berichten, 92(4506)*, 168-170.

De Witte, K. & Marques, R. C. (2010). Designing performance incentives, an international benchmark study in the water sector. *CEJOR, 18*, 189-220.

Dijkgraaf, E., de Jong, R., van de Mortel, E.G., Nentjes, A., Varkevisser, M., Wiersma, D.(1997). *Mogelijkheden tot marktwerking in de Nederlandse watersector – Possibilities for competition in the Dutch water sector*, Erasmus Universiteit Rotterdam OCFEB, Rotterdam.

Dijkgraaf, E., Mendys, E. & Varkevisser, M. (2007). *Quick scan prestaties drinkwaterbedrijven in België, Engeland en Nederland – Quick scan of performances drinking water utilities Belgium, England and the Netherlands*, SEOR-ECRi, Rotterdam.

Dijkgraaf, E., van der Geest, S. A. & Varkevisser, M. (2005). Efficiëntie boven water – Efficiency over water. *Economische Statistische Berichten, 90(4451)*, 34-36.

Dixon, J. & Kouzmin, A. (1994). The commercialization of the Australian public sector: competence, elitism or default in management education. *International Journal of Public Sector Management, 7(6)*, 52-73.

Dumaij, A. C. M. & van Heezik, A. A. S. (2012). *Productiviteitstrends in de drinkwatersector - Een empirisch onderzoek naar het effect van regulering op de productiviteitsontwikkeling tussen 1985 en 2010 – Productivity trends in the drinking water sector – An empirical study about the effect of regulation on the development of the productivity between 1985 and 2010*, TU Delft, Delft.

Goodhart, C. A. E. (1975). Problems of monetary management: the UK experience in papers in monetary economics. *Monetary Economics, 1*.

Henry, G. T. & Mark, M. M. (2003). Beyond use: understanding evaluation's influence on attitudes and actions. *American Journal of Evaluation, 24(3)*, 293-314.

Huisman, P. (2002). How the Netherlands finance public water management, *European Water Management Online*, official publication of the European Water Association (EWA), 2002(3).

ILT (2012). *Protocol prestatievergelijking drinkwaterbedrijven 2012 – Protocol benchmarking drinking water utilities 2012*, Inspectie Leefomgevinng en Transport, Den Haag.

Kouzmin, A., Löffler, E., Klages, H. & Korac-Kakabadse, N. (1999). Benchmarking and performance measurement in public sectors: Towards learning for agency effectiveness. *International Journal of Public Sector Management, 12(2)*, 121-144.

Lawrenz, F., Gullickson, A. & Toal, S. (2007). Dissemination - Handmaiden to evaluation use. *American Journal of Evaluation, 28(3)*, 275-289.

Magd, H. & Curry, A. (2003). Benchmarking: achieving best value in public-sector organizations. *Benchmarking: An International Journal, 10(3)*, 261-286.

Meyer, M. W. & Gupta, V. (1994). The performance paradox. *Research in Organizational Behavior, 16,* 309-369.

Neunteufel, R., Perfler, R., Mayr, E., Theuretzbacher-Fritz, H. & Kölbl, J. (2010). Benefits from benchmarking – an Austrian case study. *Water Practice and Technology*, *5(2)*.

Nutt, P. & Backoff, R. (1993). Organizational publicness and its implications for strategic management. *Journal of Public Administration Research and Theory, 3*, 209-240.

Osborne, S. P. (2006). The new public governance? *Public Management Review, 8(3)*, 377-387.

Png, I. (2002). *Managerial Economics*, 2. Blackwell Publishing, Malden.

PWC, (2008). *Eerst water, de prijs komt later? - Water first, the price second?* retrieved on December 2012 at http://www.vewin.nl/SiteCollectionDocuments/Nieuws%202008/Rapportage%20PWC%20waterenquete.pdf

Schouten, M. (2009). *Strategy and performance of water supply and sanitation providers. Effects of two decades of neo-liberalism*. PhD thesis, Erasmus Universiteit Rotterdam, Rotterdam.

Schmitz, T. & Dane, P. (2008). A sharp improvement in the efficiency of Dutch water utilities: benchmarking of water supply in the Netherlands, 1997-2007. *Water Utility Management International, 19, 17-19*.

Segrave, A. J. (2014). *Time to change - The foreseeable future for water planning*. PhD thesis, TU Delft, Delft.

Smith, A. (1976). *The Wealth of Nations*. Penguin Books, London.

Steenhuisen, B. M. (2012). Virtueel vangnet of functionele beleidsdroom - Publiek aandeelhouderschap als toezicht op de drinkwatersector – A virtual safety net or a functional policy dream – public ownership of shares as regulation of the drinking water sector. *Bestuurskunde 2012(3)*.

Trosa, S. & Williams, S. (1996). Benchmarking in public sector management. In: *Performance Management in Government*, D. Shand (ed.), Public management - occasional papers.

van Damme, E. E. C. & Mulder, K. J. (2006). Transparant en eerlijk geprijsd water – Transparent and faire priced water. *Economische Statistische Berichten, 91(4482)*, 134-136.

van Geel, P. L. B. A. (2006). *Nieuwe bepalingen met betrekking tot de productie en distributie van drinkwater en de organisatie van de openbare drinkwatervoorziening (Drinkwaterwet) – New definitions with regard to the production and distribution of drinking water and the organization of the public drinking water supply sector (Drinking water law)*. In: T.K.d. Staten-Generaal, 30 895. Sdu Uitgevers, Den Haag.

van Helden, G. J. & Tillema, S. (2005). In search for a benchmarking theory for the public sector. *Financial Accountability & Management, 21(3)*, 337-361.

van Helden, J. & Brouwer, N. (2005). Benchmarking en prestatieverbetering in de publieke sector - Ervaringen met de bedrijfsvergelijking zuiveringsbeheer. *Maandblad voor Accountancy en Bedrijfseconomie, 79(11),* 573-581.

Van Peursem, K. A., Pratt, M. J. & Lawrence, S. R. (1995). Health management performance - A review of measures and indicators. *Accounting, Auditing & Accountability Journal, 8(5)*, 34-70.

VEWIN (2000). *Water in Zicht 1999 - Bedrijfsvergelijking in de Drinkwatersector - Reflections on Performance - Benchmarking in the Dutch drinking water industry*. Rijswijk.

VEWIN (2004), *Water in zicht 2003 – Bedrijfsvergelijking in de Drinkwatersector - Reflections on Performance - Benchmarking in the Dutch drinking water industry*. Rijswijk.

VEWIN (2007), *Water in zicht 2006 – Bedrijfsvergelijking in de Drinkwater sector - Reflections on Performance - Benchmarking in the Dutch drinking water industry*. Rijswijk.

VEWIN (2010), *Water in zicht 2009 – Bedrijfsvergelijking in de Drinkwatersector - Reflections on Performance - Benchmarking in the Dutch drinking water industry*. Rijswijk.

VEWIN (2013). *Water in zicht 2013 – Bedrijfsvergelijking in de Drinkwatersector - Reflections on Performance - Benchmarking in the Dutch drinking water industry*. Rijswijk.

Wynn-Williams, K. L. H. (2005). Performance assessment and benchmarking in the public sector. *Benchmarking: An International Journal, 12(5)*, 482-492.