On the Impact of the Dutch Educational Supervision Act

Ehren, Melanie C. M.; Leeuw, Frans L.; Scheerens, Jaap

Published in:
American Journal of Evaluation

Publication date:
2001

On the Impact of the Dutch Educational Supervision Act

Analyzing Assumptions Concerning the Inspection of Primary Education

Melanie C. M. Ehren
University of Twente

Frans L. Leeuw
University of Utrecht and Ministry of Justice

Jaap Scheerens
University of Twente

Abstract: This article uses a policy scientific approach to reconstruct assumptions underlying the Dutch Educational Supervision Act. We show an example of how to reconstruct and evaluate a program theory that is based on legislation of inspection. The assumptions explain how inspection leads to school improvement. Evaluation of these assumptions is used to predict the (in)effectiveness of this legislation. The article concludes by discussing the advantages and drawbacks of this kind of approach as a starting point for impact and effect studies. As the program theory of inspection includes elements common to other kinds of educational interventions and reforms, these elements can also be considered for other types of program theory.

Keywords: *program theory; school inspections; supervision act; education*

Every year, the central government of the Netherlands fulfills its constitutional duty of supervising education by dispatching a team of 72 inspectors to make 2,265 visits to primary schools. These inspectors visit schools half a day every year for a short yearly inspection and 2 days every 4 years for a more elaborate quality control. This extended visit is also carried out if the yearly visit shows defects in school quality. Before this visit, schools receive a letter requesting them to send information to the inspectorate, such as their prospectus and school plan. Schools are also invited to fill in questionnaires about, for example, their pedagogical vision, their lesson tables, and the didactic they use. During the (extended) school visit, the inspector observes a number of lessons and interviews teachers, the school director, parents, and pupils. These observations and interviews are used to obtain a picture of how the school is

Melanie C. M. Ehren, University of Twente, Faculty of Behavioral Science, P.O. Box 217, 7500 AE Enschede, the Netherlands; phone: +31 (53) 489-2451; e-mail: m.c.m.ehren@edte.utwente.nl.

American Journal of Evaluation, Vol. 26 No. 1, March 2005 60-76
DOI: 10.1177/1098214004273182

doing on the standards of the inspectorate that are part of a framework for inspection. This framework contains, among others, a general description of the type of pedagogical climate the school should have, the teaching and learning strategies that should be used, and how schools should take care of pupils with learning difficulties. The framework exists of both legal requirements, which are formulated in the educational acts, and quality standards, which are formulated in the Educational Supervision Act. At the end of the visit, the inspector feeds back his or her judgment to the school and writes a school report that is published on the Internet.

Implemented in January 2003, the Educational Supervision Act specified these (and other) working methods and described the framework for inspection. The act specifies certain expectations about how schools should be inspected, the effects such inspections are expected to have, and how these effects should be realized. Together, these assumptions form the program theory of the Educational Supervision Act. Identifying and explicating this program theory enables a critical evaluation of the validity of the assumptions. In this article, we will reconstruct the program theory by specifying the intentions of the legal act. The accuracy of the assumptions will be evaluated, and the benefits and pitfalls of a program theory method for evaluation studies will be discussed.

Method

The first step is to reconstruct the assumptions that explain how inspection is supposed to work. This conglomerate of assumptions forms the program theory (Chen, 1990). Next, the reconstruction of the program theory is validated to be sure of having a legitimate overview of the assumptions. The last phase consists of a critical evaluation.

Choices to Be Made

The reconstruction phase starts with a number of choices that have to be made. First of all, the aims of reconstructing the program theory have to be stated. According to Fleurke and Huizenga (1988), possible aims may be to describe, to explain, or to predict the (in)effectiveness of a program or to analyze the way policy is developed. The program theory in this study is part of a larger research project designed to measure effects and side effects of inspection in the Netherlands. The aim of reconstructing the program theory is therefore to predict the (in)effectiveness of inspection by describing and evaluating how schools should be inspected according to the Educational Supervision Act, the effects such inspections should have, and how these effects should be realized.

Second, the method of gathering information about assumptions should be chosen. The (re)construction of the program theory should be explained, and the way in which assumptions are explained should be described. As the Educational Supervision Act is leading in our study, analyzing this act and the documents that are related to it is a logical starting point. Other possibilities are group discussions and interviews with relevant respondents. The advantage of documents in relation to interviews and discussions is, however, that expressions are often more consistent (Karstanje, 1996). Documents such as the memorandum of explanation and the minutes of governmental discussion concerning the act are also vital for interpreting the act when it is challenged in a legal setting, such as a courtroom.

Translating the information into assumptions can be done in a number of ways. Selection of one of these methods is the third choice to be made. Leeuw (2003, p. 7) described a number of methods (that relate to the way information is gathered). We will use the policy scientific approach, which consists of the following steps:

1. *Identify the social and behavioral mechanisms that are expected to solve the problem*; search formal and informal documents for statements indicating the necessity of solving the social, organizational, or policy problem in question, the goals of the proposed policy or program, and how they are to be achieved. These latter statements refer to mechanisms (or "engines") that drive the policies or programs and are believed to make them effective. Examples are manifold. They include determinants of innovation diffusion, mechanisms underlying Prisoner's Dilemma games, processes producing social capital, cognitive dissonance, different types of learning behavior, and many more. Statements having the following form are especially relevant for detecting these mechanisms:
   - It is evident that *x* will work.
   - In our opinion, the best way to address this problem is to . . .
   - The only way to solve this problem is to . . .
   - Our institution's *x* years of experience tell us that . . .
2. *Compile a survey of these statements and link the mechanisms to the goals of the program under review.*
3. *Reformulate the statements into conditional "if-then" propositions or propositions of a similar structure* (e.g., "the more *x*, the less *y*").
4. *Search for warrants that will identify disconnects in or among different propositions using argumentation analysis*. Founded in part on Toulmin's (1964) *The Use of Argument*, argumentation analysis refers to a model for analyzing chains of arguments and helps to reconstruct and "fill in" argumentations. A central concept is the warrant, which, according to Toulmin (1958) and Mason and Mitroff (1981), is the "because" part of an argument. A warrant says that B follows from A because of a (generally) accepted principle. For example, "the organization's performance will not improve next year" follows from "the performance of this organization has not improved during the past 5 years" because of the principle that past performance is the best predictor of future performance. The "because" part of such an argument is often left implicit, with the consequence that warrants must be inferred by the person performing the analysis.
5. *Reformulate these warrants in terms of conditional "if-then" (or similar) propositions and draw a chart of the (mostly causal) links*.

Examples

To clarify these steps, two examples will be given of how statements were used to reconstruct assumptions. These examples primarily explain the first three steps. The first example clarifies how a single statement, containing a global description, is rephrased into a hypothetical if-then assumption. As the reconstruction only calls for a sharpening, this is an easy reconstruction.

Example 1

The statement "Compliance control will focus on ensuring a certain level of quality"[1] translates into the following assumption: "If the inspectorate controls compliance of schools, then schools will (eventually) attain a satisfactory level of quality."

The second level of interpretation concerns matching different statements to form a single assumption.

Example 2

The statements "The inspectorate should stimulate schools' responsibility for enhancing the quality of education"; "as a learning organization, a school should have primary responsibility

for monitoring and improving its own quality;"[2] and "When they are rewarded with less intense and less frequent inspection, schools will be motivated to develop good quality assurance systems"[3] translate into the following assumption: "If the inspectorate decreases the intensity and frequency of inspection in case a school has developed a quality assurance system and delivers good educational quality, schools will develop systems for quality assurance."

Participant Check

After reconstructing the program theory, the reconstruction is validated, in the sense of being reviewed by relevant actors, to be sure that the assumptions actually represent the intentions of the act. This kind of participant check is often described as a criterion the reconstruction should meet. Performing it separately enables us, however, to be more specific about faulty interpretations during the reconstruction phase. The intermediate adjustments can be explained and accounted for.

A first list of reconstructed assumptions was checked by interviewing 14 people responsible for the original statements. Members of the House of Representatives of the States General, employees of the inspectorate and the Ministry of Education, and members of the Educational Council were asked to indicate whether the assumptions (translated into statements) fit the intention of the Educational Supervision Act. If not, interviewees were asked why, in their view, this was not the case.

This participant check led to combining two forms of inspection (inspection of quality standards, which are described in the Educational Supervision Act, and inspection of compliance with legal requirements, which are formulated in the educational acts) into one (inspection of school quality). No other adjustments were made. The Results section shows the assumptions that correspond to the intentions and assumptions of those involved in creating the act.

Evaluation

A critical evaluation of the program theory is the last phase of the study. Results of prior research are used to analyze how consistent, complete, and realistic the assumptions are. By doing so, the potential effects (and side effects) of the act can be predicted. If prior research shows that the assumptions have a low potential of actually meeting the desired effects, a number of conclusions may follow. One of them can be to adjust the act or to use other variables in further investigation of effects from educational inspections.

Results

Our starting point for reconstructing the assumptions was the memorandum *Varieteit en Waarborg* (Variety and Guarantee; Ministry of Education, Culture and Science, 1999), in which the Minister of Education sketched the basics of renewed inspection. The publication of the Educational Supervision Act in the *Government Gazette*, 2.5 years later, forms the beginning of this renewed inspection.

The first memorandum (p. 8) specifies two functions of supervision: (a) Through inspection, the government guarantees that schools will deliver a satisfactory level of educational quality for all citizens, and (b) through inspection, the government stimulates schools to develop their own quality assurance system, which will lead to improvement in the quality of education. The discussion about the Educational Supervision Act focused on how these functions should be carried out. The results in this section convey the final choices that were made.

According to the Educational Supervision Act, guaranteeing a satisfactory level of educational quality is achieved primarily through inspection of legal requirements that schools should comply with. These requirements are placed on schools by educational laws; they are a prerequisite for receiving governmental funding.[4] Because the legal requirements alone are considered to be insufficient for achieving the goal of improved education, the Educational Supervision Act further specifies quality standards (that partly elaborate on the legal requirements). The inspectorate is entitled to evaluate these quality standards in such a manner that schools are challenged to provide higher standards of quality than those considered merely satisfactory.[5] *Challenged* is the key word here as schools are not obliged to meet the quality standards, as they are obliged to comply with legal requirements. The quality standards are only intended to guide the working methods of the inspectorate, which consist of assessing the schools' performance and confronting the schools with their strong and weak points in relation to a framework for inspection. As a matter of fact, this framework contains both the legal requirements that schools should comply with and the quality standards that inspectors use to assess schools. Such confrontation of schools with their strong and weak points is expected to make schools that are performing below standards aware that they can and should do better.[6] As schools are seen as organizations that are willing to change, information about weak points is considered to be a major challenge and stimulant for change.

The Educational Supervision Act also provides for proportional inspection, in which the amount and frequency of inspection varies according to the results of quality assurance of schools.[7] Schools that have a quality assurance system that provides reliable information about the inspection standards are visited less frequently and intensively. The inspectorate uses this information for its own school reports and does not have to gather all the necessary information itself by visiting schools.

The publication of the inspection reports is the final method employed by the inspectorate to stimulate quality.[8] Parents are expected to take note of the report and act upon it. These actions should also challenge schools to change.

It is assumed in the Educational Supervision Act that these forms of inspection are capable of generating improvements in the quality of education. Quality is described as the provision of equal educational opportunities to all young people[9] and as the value added by schools in terms of student achievement.[10] Three main assumptions that may be articulated from this description of the pillars of the program theory reconstructed in our study are the following:

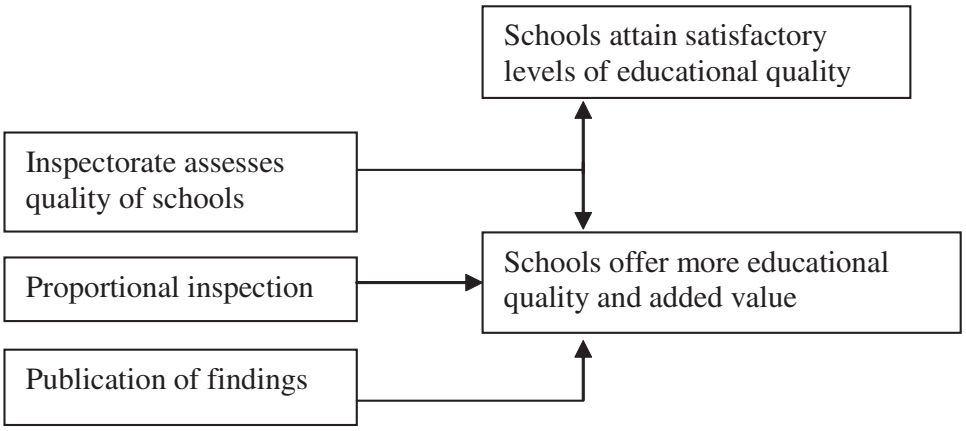
- If the inspectorate assesses the quality of schools, schools will attain satisfactory levels of quality and will also offer more added value in terms of student achievement.
- If the inspectorate employs proportional inspection, schools will offer more added value.
- If the inspectorate publishes its findings in an accessible manner, schools will offer more added value.

The model presented in Figure 1 shows these assumptions, which will be described in more detail below. The subsequent paragraphs will elaborate on this framework.

Quality Assessment

The framework used by the inspectorate to assess school quality integrates two types of indicators: legal requirements and quality standards.[11] Legal requirements are formulated in the separate educational acts. Schools have to meet these requirements to receive governmental financing. These requirements describe, for example, a number of educational goals or state that schools should prepare their pupils for secondary education. Quality standards, such as an

Figure 1
Framework for Program Theory



adequate pedagogical climate and supply of subject matter, are part of the Educational Supervision Act and are intended to guide the inspectorate in assessing schools. Schools are not obliged to meet these criteria.

As legal requirements and quality standards overlap to a considerable extent[12] and quality standards often refine the legal requirements, they are integrated in the framework for inspection. However, freedom of education is highly valued in the Netherlands and implies that schools are free to determine how they will organize their education. As long as they comply with legal requirements, they cannot be sanctioned or be obliged to change.[13] The official documentation[14] of the debate regarding the Educational Supervision Act emphasizes that the inspectorate should distinctively inform schools about failing to comply with legal requirements or failing to meet the quality standards. In general, policy makers expect schools to support the framework, as it was developed in close cooperation with educational organizations that represent schools.[15] Schools that do so will consequently also use the feedback of the inspectorate, regardless of whether that feedback concerns legal requirements or quality standards.[16]

How schools should use the feedback, and what role the inspectorate should take in this process, is not clear. On one hand, the program theory emphasizes the school's own responsibility for improvement, whereas on the other hand, it argues that schools should be guided by inspection.

The first argument, about schools carrying responsibilities for their own quality, stresses that each school should consider its environment in deciding how to improve and act in the best interest of its pupil population.[17] The inspectorate provides schools with information about their strengths and weaknesses[18] and asks schools to account for the choices they have made instead of making these choices for them.

For schools considered qualitatively weak in ways that go beyond meeting legal demands, a different approach is indicated, including more elaborate visits, possible extra financial means, and the requirement to develop plans for improvement that should be assessed by the inspectorate.[19] In this case, over and above its evaluative role, the inspectorate should also fulfill an advisory function for weaker schools, specifying possible improvements and explaining how things

can be done differently.[20] The documents refrain from giving a definition of *weak schools*. The Memorandum of Explanation (p. 6) specifies, however, that the improvement trajectories of schools are explicitly voluntary, unless schools fail to comply with legal requirements. The advisory function therefore seems to be applied to schools that fail to comply with legal requirements.

The reconstructed program theory states the expectation that school improvement will always lead to higher quality in educational processes. The improvement of educational processes (or educational quality), in its turn, leads schools to offer satisfactory levels of educational quality and to add more value in terms of student achievement.[21]

Added value describes the result of a process in which schools bring their pupils up to the level of performance that can be expected of them—taking into account the students' performance levels at entry—while minimizing the number of students who either leave school or repeat classes.[22]

Table 1 provides an example of how the assumptions, concerning the quality assessment of the inspectorate, can be described in a more explicit and scheduled way.

Proportional Inspection

The proportional working method of inspection describes the use of results of quality assurance and self-evaluation of schools by the inspectorate to form judgments about school quality.[23] To be helpful, the self-evaluation results of schools must be reliable and provide information about the indicators included in the inspection framework.[24] If these requirements are met, the inspectorate will confront the schools with fewer and less intense inspection visits.[25]

However, the Educational Supervision Act also provides for annual visits. These are considered to be necessary for the frequent monitoring of schools.[26] To maintain the proportional character of inspection, this annual visit should focus on the smallest possible selection of standards (including quality assurance of schools and pupil learning results, among others) necessary to assess the functioning of the schools.[27]

The working methods and the capacity of the inspectorate are two ways in which proportional inspection is expected to increase the value added by schools. First, the proportional inspection method is expected to stimulate schools to develop quality assurance measures and to evaluate their performance according to their own standards.[28] A prerequisite for proportional inspection is that schools are aware of, and agree with, this method of work.[29] By consulting with educational organizations that represent schools about the framework for inspection, the inspectorate strives to ensure that schools both agree with and support the proportional working method.[30]

As a result of their quality assurance efforts, schools are expected to be capable of identifying and correcting their own weak points. The documents state that schools are expected to develop quality assurance that is more elaborate than the conditions specified by the inspectorate for proportional inspection. These conditions state that the inspectorate uses the results of quality assurance of schools if they provide reliable information about the indicators in the inspection framework. However, schools should not limit their focus to these conditions but should develop and assess their own quality norms in close contact with, and adjusted to, their pupil population and environment.[31]

Quality assurance is expected to lead to increased autonomy for schools to make their own choices concerning quality policy and to adjust their education to their own context and pupil population.[32] Schools that adjust their education to pupils will eventually add more value.[33]

Table 1
Assumptions About Quality Assessment

1. If the inspectorate assesses school quality, schools will attain satisfactory levels of educational quality and offer more added value.
   - 1.1 If the inspectorate assesses school quality, it will do so using a framework that provides insight into the value added by schools and into the quality standards necessary to offering added value.
   - 1.2 If the inspection framework provides insight into the value added by schools and into the quality standards necessary to offering added value, this framework will be found both realistic and relevant by schools.
     Because the framework is developed in close cooperation with umbrella organizations in the educational field, which ensures that the framework is feasible for schools.
     Because schools will support a framework that is developed in close cooperation with the educational field.
   - 1.3 If schools find the framework realistic and relevant, they will provide the inspectorate with all of the information that is needed for carrying out the inspection.
   - 1.4 If all relevant information is available, the inspectorate can provide feedback on the strengths and weaknesses of the schools.

For strong schools:

   - 1.5 If the inspectorate provides schools and other stakeholders with feedback concerning the strengths and weaknesses of particular schools (in dialogue, by separating legal requirements and quality standards, mentioning only the areas in which particular schools should improve, but not how to do so), schools will improve autonomously.
   - 1.6 If schools improve, they will yield more added value.
     - 1.6.1 Autonomous improvement processes are expected to ensure that education is adjusted to the needs of pupils.
       Provided that governmental regulations and inspection give schools enough leeway to do so.
     - 1.6.2 If schools adjust education to the needs of their pupils, more added value will be offered to all segments of the pupil population, regardless of initial achievement levels.

For weak schools:

   - 1.7 If the inspectorate provides schools and other stakeholders with feedback about strengths and weaknesses, schools will use this feedback to improve.
     - 1.7.1 If the inspectorate provides schools and other stakeholders with feedback about strengths and weaknesses, the inspectorate will also outline the measures to be taken (e.g., more visits) and how schools can improve.
     - 1.7.2 If the measures to be taken by the inspectorate and the guidelines for improvement are clear, schools will improve on the basis of the feedback obtained from the inspectorate.
       Because the inspection framework is based on research and legitimated through consultation with educational organizations that represent schools.
   - 1.8 If schools improve as a result of using feedback from the inspectorate, educational quality will be enhanced.
   - 1.9 If educational quality improves, schools will offer more added value in terms of student achievement.

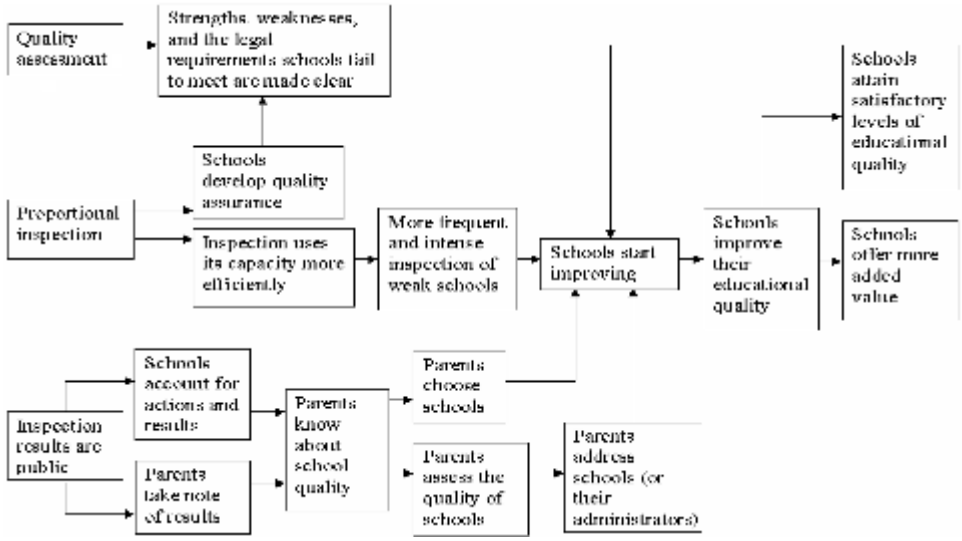
The second way in which proportional inspection leads to increased added value is through efficient allocation of inspection capacity. The intensity with which weak schools are inspected is assumed to increase, whereas inspection capacity stays the same.[34] For this to happen, however, the inspectorate must be capable of identifying these schools by assessing the soundness of procedures and the actual results of their self-evaluations, that is, the assessments schools make of their own educational quality in relation to the inspection framework.[35]

If a weak school is subjected to more frequent and intense inspections, it is assumed that the school will improve and will eventually offer more added value and, consequently, a satisfactory level of educational quality. This assumption is based on the fact that the inspection of weaker schools is linked to the requirement that schools present improvement plans that will be closely monitored by the inspectorate.[36]

Publication of Findings

The findings of school inspections are published with the intent of giving the environment of the school a role in school improvement. In particular, parents should be able to participate,

Figure 2
Program Theory Concerning the Educational Supervision Act



using the independent, accessible, and public information provided by the inspectorate concerning schools. The expectation is that parents, more than any other stakeholders, will use this information to evaluate the quality of their children's schools.[37] If a school does not meet their expectations for quality, it is assumed that parents will address the school (or its administrators) about possible improvements,[38] thereby leading schools to improve and add more value.[39]

Parents can also participate in school improvement by choosing schools that offer high educational quality; the public information provided by the inspectorate can help them in this selection process.[40] If schools are aware that good schools are popular, they will be stimulated to improve.[41]

Public information about schools should stimulate schools to explain their educational processes and the learning results of their pupils. As a result, schools will be accountable to their environments for their actions (and for the results of these actions).[42] Parents will thus be able to inform themselves about the quality of their children's schools and to address the schools if their expectations for educational quality are not met.

The program theory assumes that inspection leads to satisfactory quality of educational processes and to more added value in terms of student achievement. To stimulate educational quality (in either way), inspection must consist of a proportional working method using a framework for inspection that results in the publication of information concerning the quality of schools. Figure 2 shows the relationships among these assumptions.

Evaluation

The program theory seeks to clarify how inspection should lead to certain effects. How consistent, complete, and realistic are the underlying assumptions? The answer to this question suggests possible flaws in the program theory that arise when assumptions are incorrect and

when inspection (as planned) is not able to produce the desired effects (Weiss, 1997). Results from prior research provide clues concerning the accuracy of the assumptions.

Consistency

The program theory is inconsistent in a number of ways. These all seem to be related to the freedom of education that is laid down in Article 23 of the Dutch Constitution. This article implies that as long as schools meet certain legal requirements, government and inspection have no formal responsibility for school improvement. One might even say that Article 23 implies that the inspectorate should stay away from pedagogical interference. This fits the current societal and political opinion in the Netherlands that schools should be autonomous and have their own responsibilities in providing and organizing their education. What the exact place and responsibilities of the inspectorate are is not clear though. In a system where government deregulates, the inspectorate is considered to be the countervailing power at the end of the chain.[43] Schools have to be in charge of their own quality but should at the end account for the actions they have taken and the results they have achieved. However, a stable and accepted modus of autonomy of schools in relation to deregulation by government has not been found yet. The tasks and responsibilities of the inspectorate that are, in any way, related to improvement of schools seem to be ambiguously devised as a result of obscurities in this system. A clear definition of how far autonomy of failing schools goes and the exact responsibilities of the inspectorate when encountering a failing school is needed. When the ultimate goal is to improve failing schools, a simple statement that the inspectorate can only name and report about failing schools does not seem to be sufficient.

A second ambiguity exists in the framework for inspection that integrates legal requirements and quality aspects. Schools are obliged to improve on the legal requirements, whereas they are free to decide on possible improvements in relation to the quality aspects. The first part in the feedback of inspectors is therefore compulsory, whereas the second part is voluntary. On one hand, the inconsistency concerns the integration of legal requirements and quality aspects in the framework for inspection, whereas on the other hand, it concerns the needed separation of these in the feedback provided by inspectors.

An inconsistency also arises after inspectors have given their feedback to schools. The program theory speaks about different forms of inspection for strong and weak schools. Weak schools receive more elaborate visits and need to develop plans for improvement that are assessed by the inspectorate. However, improvement trajectories are voluntary unless schools do not meet the legal requirements. As the legal requirements are not assessed separately but are part of a more elaborated framework to assess school quality, it is not logical to define weak schools by just referring to the legal requirements. It may therefore happen that the improvement trajectories are assigned to a larger contingent of schools than just the ones that fail to meet the legal requirements. Some schools run the risk of being confronted with an improvement trajectory that is presented as compulsory, whereas, in fact, it is not. Other schools may not be confronted with an improvement trajectory at all, whereas, in fact, they should. Again, the extent to which schools are free to decide for themselves whether they need to improve and how autonomous they are may not be clear in reality.

A last inconsistency concerns the proportional inspection. Schools are expected to develop their own quality assurance system and to use their own quality norms in close contact with their pupil population and environment, whereas, in fact, they are rewarded for a specific quality assurance system that only gives reliable information about the indicators in the inspection framework. Another example of how rewarding may lead to unwanted results is when weak

schools receive extra financial means to improve, whereas strong schools do not receive extra money.

Completeness

The program theory lacks a number of definitions that are necessary to implement the Educational Supervision Act correctly or to assess if it is meeting the results that it intents to meet. The assumptions fail to explain, for example, what "a satisfactory level of educational quality" is. In connection to this incompleteness, a distinction between strong and weak schools and an allocation of an *evaluative* or an *advisory* inspection to strong and weak schools is hard to make.

The program theory is also unclear about, for example, the competencies of inspectors or the way schools should account for their results. We will not elaborate on these factors here, as they are considered to be factors that are not regulated in legislation and are therefore not part of this program theory.

Realism

The program theory assumes that schools use the feedback of inspection to improve. Improvement is expected to lead to satisfactory levels of educational quality and more added value. Schools are assumed to support both inspection and feedback from inspection and are assumed capable of improvement (with some support from their environment). The program theory also provides for the publication by the inspectorate of the results of school inspections. Parents are assumed to stimulate schools to improve by recommending possible improvements and by choosing only high-quality schools. Some (likely) flaws in the theory are considered below.

Feedback provided by the inspectorate or resulting from quality assurance does not necessarily lead to improvement. There are a number of prerequisites for feedback to lead to improvement (Ilgen, Fisher, & Taylor, 1979; Kluger & DeNisi, 1996). First of all, the school needs to experience the feedback as relevant, understandable, clear, and useful (Brimblecombe, Shaw, & Ormston, 1996; Doolaard & Karstanje, 2001). Feedback of the inspectorate has a larger chance of being used when teachers are involved in recommendations, when support is given to the school (Gray & Wilcox, 1995), when schools have insight in their own strong and weak points, when their culture is open to feedback and outside criticism (Early, 1998), and when inspectors illustrate both the cause of bad performance as well as its remedy (Doolaard & Karstanje, 2001; Gray & Gardner, 1999).

Even when schools do use the feedback provided by the inspectorate to start improving, it does not necessarily lead to the desired results. According to Geijsel (2001) and Reezigt (2001), the likelihood that a school will successfully improve depends on such internal features as cooperation between teachers and organizational learning, as well as on the context of the school (e.g., the level of support provided by parents).

Prior research shows that schools that start to change after inspection mainly focus on enhancing commitment of teachers to the school, formulating behavioral rules for pupils, setting up strategies for improving examination results, and changing management style and structure (Kogan & Maden, 1999, p. 18; Visscher, 2002, p. 62). Inspection does not influence educational and subject goals, nor the didactic, methods, division of labor, or internal communication in the school (Brimblecombe et al., 1996; Chapman, 2001). Research in the United Kingdom

even shows a decline in achievement levels in the year of the inspection visit (Cullingford, Daniels, & Brown, 1998; Rosenthal, 2004; Shaw, Newton, Aitkin, & Darnell, 2003).

Parents do not use the public information about schools provided by the inspectorate. Research shows that parents are interested in matters other than inspection results (e.g., Dronkers & Veenstra, 2001; Educational Council, 2001; Karsten & Visscher, 2001). They are primarily concerned about the atmosphere, pedagogical climate, working methods, safety, clarity of regulations, waiting lists for special education, reputation of the school, and about decisions concerning the promotion of pupils to the next class. When parents do suggest possible improvements, these usually involve such conditional matters as timetables (Educational Council, 2001). Parents do not generally interfere in matters relating to educational quality, nor are their school choices based on educational quality. Only 2% of the parents (mostly those who are more highly educated) use pupil learning results as a criterion for choosing a school (Dronkers & Veenstra, 2001, p. 33). For most parents, proximity, cost, and religious foundation play a larger role in school choice. According to Hargreaves (1995), there is also little evidence of parents seeking alternatives to an exposed failing school.

Perhaps this will change when public information about schools on the Internet is more commonly used and available to more parents than just the higher educated ones, but for now, we cannot expect a large number of activities from parents.

Schools are not inclined to use criticism from parents for school improvement. Suggestions from parents for possible improvements do not usually result in actual improvement. Because parents do not pay the bill for education, schools do not feel that it is necessary to take preferences of parents into consideration (Educational Council, 2001). In the Netherlands, the present system of educational financing gives schools very little incentive to perform better, as weak schools are ensured of enough pupils and weak teachers maintain their salaries (Waterreus, 2003, p. 33). When schools must compete for pupils, they are likely to be more inclined to improve their public relations and after-school activities than they are to improve their educational quality.

Possible side effects are left out. Schools do not always have an interest in cooperating during the inspection visit or to perform all the activities that are necessary to improve after or in relation to inspection. A school, for example, must provide information on (painful) matters so that inspectors can identify and publish the weak spots. According to Kerr (1975), people will more readily perform behaviors for which they will be rewarded. In this case, schools clearly have no incentive for revealing their weak points to the inspectorate. Schools are much more likely to anticipate the inspection visits and to behave in different ways as they usually do (Smith, 1995, p. 280). Chapman (2001) found, for example, that teachers prepare and structure their lessons better when inspectors visit the school. Some schools even take improper actions, according to Wiebes (1998), such as sending risky pupils to another school, selecting pupils before entrance, withdrawing from or training pupils for tests, or having pupils stay down a class. The public results of inspection are a large cause for this kind of behavior, according to the Educational Council (2001) and Dijkstra, Karsten, Veenstra, and Visscher (2001).

Smith (1995) also found that schools might focus on the indicators in the framework when planning their educational processes. This can lead to ossification, that is, schools may refrain from undertaking innovative activities that do not fit the framework and are thus not rewarded by the inspection system.

Although it is too early to present any results from proportional inspection, the same type of argumentation can be made for this working method. The inspectorate uses self-evaluation

results only if they meet strict conditions (they must provide reliable information about the indicators in the inspection framework). As a result, schools are not rewarded for quality assurance systems using indicators that differ from the inspection framework.

The situation described above could result in schools starting to look alike, at least in terms of inspection findings (Dimaggio & Powell, 1983). The performance indicators in the framework for inspection lose their sensitivity in detecting bad performance or low quality. Schools learn which aspects of their performance are measured and which are not. They can use this knowledge to influence the assessment of the inspectorate, mainly by investing in improving the performance that will be measured rather than aiming for underlying goals. This "perverse learning" (Van Thiel & Leeuw, 2003) occurs, for example, when schools invest in developing and implementing a certain type of pupil monitoring system, described in the framework as a good example, instead of focusing on how to give pupils the best possible care.

Discussion

In this article, we reconstructed the program theory of the Dutch Educational Supervision Act to describe how inspections should lead to certain results. The merit of this method lies in the possibility to give a critical evaluation of assumptions behind innovations or interventions; in this study, the school inspections as laid down in the Educational Supervision Act. A second merit is the link between theory and practice that is made when reconstructing program theory as a starting point for further research. The means, goals, and intervening mechanisms can be used to develop variables for further investigation. If the results of further investigation show no effects, these can be explained by comparing real program implementation to intended implementation. If these do not match, an explanation for the absence of effects may be faulty program implementation.

A program theory approach has some drawbacks, however. Perrin (1999) stated a major one by saying that it is not sufficient to judge the value or impact of a program or initiative based only on its intent and on what it claims to do. It is essential to examine what really happens and the extent to which the outcomes are a result of the program intervention or are due to other factors. Elte (1988, p. 2) also underlined the importance of looking at real (implemented) means and goals by saying that informal assumptions mostly influence the functioning and results of organizations instead of formal intentions.

Another drawback of this kind of program theory can also be the prescription to reconstruct program statements in causal terms. In reality, legislation is not always based on causal links. The method of analyzing documents in search of statements that link means to goals may even do harm to real intentions and assumptions behind legislation or may result in a model that does not correspond to reality. Some authors (e.g., Hoogerwerf, 1984) seem to solve this problem by adding other types of relations such as the ones between causes and consequences/effects and between principles and values or between principles and values on the one side and existing or expected situations. We attempted to solve this problem by organizing a separate participant check in which policy makers assessed the adequacy of the assumptions in relation to the intentions of the act.

Using a program theory approach for further research is particularly useful when the objective is to try and find possible implementation failures that occur when a program is not implemented as intended. One should be careful, however, in using a program theory approach for investigating effects of interventions, such as school inspections, when the theoretical evaluation shows that the described assumptions are not consistent, complete, and realistic. After all, it does not seem meaningful to use an incorrect implemented program theory for further investi-

gation, as empirical research shows that an incorrect program theory can lead to ineffective policy (Walraven, 1991, p. 12).

A correct, realistic, and implemented program theory, on the other hand, also does not guarantee success as other factors may play a role, such as competencies of inspectors and the capacity of schools to change. However, a correct and realistic program theory can be a good framework for searching for and filling in these other factors. This may be especially important when no control group or zero point of measurement exists and experimental research is out of the question.

This study has, so far, only shown a number of theoretical failures. The actual implementation of the Educational Supervision Act may correct these. Further research is needed to investigate whether this is indeed the case and whether this means that inspection leads to the desired results.

Notes

1. Minutes of the Senate, number 27783, 186b, p. 13.
2. Minutes of the House of Representatives of the States General, number 26572, 7, p. 7.
3. Minutes of the House of Representatives of the States General, number 27783, 21, p. 17.
4. Minutes of the House of Representatives of the States General, number 26572, 3, p. 11.
5. Minutes of the House of Representatives of the States General, number 27783, 21, pp. 24-25 and number 26572, 3, p. 4.
6. Memorandum "Variety and Guarantee," 1999, p. 9.
7. Minutes of the House of Representatives of the States General, number 27783, 21, p. 17.
8. Minutes of the House of Representatives of the States General, number 27783, 5, p. 6. Minutes of the Senate, number 27783, 32, p. 1547.
9. Minutes of the House of Representatives of the States General, number 27783, 5, pp. 17-18.
10. Minutes of the House of Representatives of the States General, number 26572, 7, p. 41.
11. Minutes of the House of Representatives of the States General, number 27783, A, p. 2.
12. Advice of the Educational Council "Deugdelijk Toezicht," 1999. Memorandum of Explanation of the Educational Supervision Act, p. 10.
13. Minutes of the House of Representatives of the States General, number 27783, A, p. 2; number 27783, 3, p. 20; number 27783, 4, p. 10.
14. Minutes of the House of Representatives of the States General, number 26572, 3, pp. 2, 7, 6, 20, 36.
15. Minutes of the House of Representatives of the States General, number 26572, 3, pp. 2, 10; number 27783, 5, pp. 2, 29, 34; number 27783, 21, p. 21. Minutes of the Senate, number 27783, 186a, p. 2.
16. Minutes of the House of Representatives of the States General, number 26572, 3, p. 3; number 27783, 3, pp. 1, 9-10. Minutes of the Senate, number 27783, 32, pp. 1554, 1566. Memorandum "Variety and Guarantee," 1999, p. 9.
17. Minutes of the House of Representatives of the States General, number 27783, 5, pp. 3, 17; number 27783, 29, p. 1; number 27783, 21, p. 22. Minutes of the Senate number 27783, 186a, p. 1; number 27783, 186b, p. 6; number 27783, 32, p. 1550. Memorandum "Variety and Guarantee," 1999, p. 5.
18. Minutes of the House of Representatives of the States General, number 27783, 5, p. 15; number 27783, 21, p. 24; number 27783, A. p. 3; number 26572, 7, pp. 7-9; number 27783, 5, pp. 3, 4, 8, 9, 15, 37; number 27783, 3, pp. 6, 29. Memorandum "Variety and Guarantee," 1999, p. 9. Minutes of the Senate, number 27783, 32, pp. 1553, 1568. Advice of the Educational Council "Deugdelijk Toezicht," 1999, p. 6.
19. Minutes of the Senate, number 27783, 186b, p. 11.
20. Minutes of the House of Representatives of the States General, number 27783, 4, pp. 13, 25-26; number 26572, 3, pp. 2, 6-7, 15; number 26572, 7, pp. 22, 26; number 27783, 22, p. 4; number 27783, 17, p. 1; number 27783, 5, p. 25; number 27783, 21, pp. 4, 27. Memorandum "Variety and Guarantee," 1999, p. 9. Minutes of the Senate, number 27783, 186a, p. 6; number 27783, 32, p. 1573.
21. Minutes of the House of Representatives of the States General, number 27783, 21, p. 31; number 26572, 7, p. 22; number 26572, 3, pp. 2, 7; number 27783, 5, p. 39. Memorandum "Variety and Guarantee," 1999, p. 15. Advice of the Educational Council "Deugdelijk Toezicht," pp. 14, 23.
22. Minutes of the House of Representatives of the States General, number 27783, 3, p. 40.
23. Minutes of the Senate, number 27783, 186b, p. 4.
24. Minutes of the House of Representatives of the States General, number 27783, 21, p. 17.

25. Minutes of the House of Representatives of the States General, number 27783, 3, pp. 6, 27; 16, 30, 15; number 27783, 4, p. 20; number 27783, 5, pp. 24, 2; number 27783, 22, pp. 3, 6; number 27783, 29, pp. 1-2. Minutes of the Senate, number 27783, 186b, pp. 4, 10, 13, 18; number 27783, 32, p. 1567.

26. Minutes of the House of Representatives of the States General, number 27783, 22, p. 1.

27. Minutes of the House of Representatives of the States General, number 27783, 22, p. 1.

28. Minutes of the House of Representatives of the States General, number 27783, 21, p. 17; number 27783, 5, pp. 22, 24; number 26572, 7, p. 21; number 27783, 3, pp. 6, 21. Minutes of the Senate, number 27783, 186b, p. 3; number 27783, 186b, p. 9. Memorandum "Variety and Guarantee," 1999, p. 14.

29. Minutes of the House of Representatives of the States General, number 27783, 5, p. 31. Memorandum "Variety and Guarantee," 1999, p. 13.

30. Memorandum of explanation, p. 13. Ministry of Education, Culture and Science, Memory of answer, April 2002, p. 3. Inspectorate of Education, framework for inspection, December 2002, p. 6. Minutes of the Senate, number 27783, 186b, pp. 4-5; number 27783, 186a, p. 2.

31. Memorandum "Variety and Guarantee," 1999, p. 13.

32. Minutes of the House of Representatives of the States General, number 27783, 5, p. 2; number 27783, 29, p. 2; number 27783, 3, pp. 2, 3, 6, 21; number 26572, 7, p. 21. Minutes of the Senate, number 27783, 186b, pp. 1, 11. Memorandum "Variety and Guarantee," 1999, pp. 2, 8-9.

33. Minutes of the Senate, number 27783, 32, p. 1550; Minutes of the House of Representatives of the States General, number 26572, 7, pp. 3, 8.

34. Minutes of the House of Representatives of the States General, number 27783, 21, p. 15; number 27783, 3, pp. 16, 30; number 27783, 5, pp. 24, 42. Minutes of the Senate, number 27783, 186b, p. 18.

35. Minutes of the House of Representatives of the States General, number 26572, 7, p. 4.

36. Minutes of the House of Representatives of the States General, number 27783, 21, p. 8. Minutes of the Senate, number 27783, 186b, p. 11.

37. Minutes of the House of Representatives of the States General, number 27783, 3, pp. 40-41; number 27783, 21, p. 22; number 26572, 3, p. 3; number 26572, 7, pp. 3, 7; number 27783, A, p. 4; number 27783, 22, p. 3; number 27783, 5, p. 3. Advice of the Educational Council "Deugdelijk Toezicht," p. 6. Minutes of the Senate, number 27783, 32, p. 1549. Memorandum "Variety and Guarantee," 1999, p. 4.

38. Minutes of the House of Representatives of the States General, number 27783, 21, pp. 13, 17, 25; number 26572, 7, p. 3; number 27783, 5, p. 6; number 27783, 3, pp. 6, 28; number 27783, A, p. 5. Advice of the Educational Council "Deugdelijk Toezicht," p. 15. Minutes of the Senate, number 27783, 186b, p. 13. Memorandum "Variety and Guarantee," 1999, p. 5.

39. Memorandum "Variety and Guarantee," 1999, pp. 13, 17. Minutes of the House of Representatives of the States General, number 26572, 3, pp. 2, 3, 5, 13; number 26572, 7, p. 4; number 27783, 3, p. 2. Minutes of the Senate, number 27783, 32, p. 1547.

40. Minutes of the House of Representatives of the States General, number 27783, 21, pp. 7-8, 13; number 26572, 7, p. 10; number 27783, A, p. 5. Memorandum "Variety and Guarantee," 1999, pp. 5, 13. Minutes of the Senate, number 27783, 186b, p. 13.

41. Minutes of the House of Representatives of the States General, number 26572, 7, p. 4; number 26572, 3, pp. 3, 5; number 27783, 3, p. 2. Minutes of the Senate, number 27783, 32, p. 1547.

42. Minutes of the House of Representatives of the States General, number 27783, 3, p. 19; number 26572, 3, pp. 11, 13. Minutes of the Senate, number 27783, 186b, p. 11.

43. Memorandum of Explanation of the Educational Supervision Act, p. 3.

References

Brimblecombe, N., Shaw, M., & Ormston, M. (1996). Teachers' intention to change practice as a result of OFSTED school inspections. *Educational Management & Administration*, *24*, 339-354.

Chapman, C. (2001). Changing classrooms through inspection. *School Leadership and Management*, *21*, 59-73.

Chen, H. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.

Cullingford, C., Daniels, S., & Brown, J. (1998). *The effects of OFSTED inspections on school performance*. Huddersfield, UK: Professor Cullingford School.

Dijkstra, A. B., Karsten, S., Veenstra, R., & Visscher, A. J. (2001). *Het oog der natie: Scholen op rapport; standaarden voor de publicatie van schoolprestaties* [The nation's eye; standards for publishing school performance]. Assen, the Netherlands: Koninklijke Van Gorcum BV.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*, 147-160.

Doolaard, S., & Karstanje, P. (2001). Gebruik van publieke prestatie-indicatoren voor schoolverbetering [Use of public performance indicators for school improvement]. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: Scholen op rapport; standaarden voor de publicatie van schoolprestaties* (pp. 155-173). Assen, the Netherlands: Koninklijke Van Gorcum BV.

Dronkers, J., & Veenstra, R. (2001). Schoolprestatie-indicatoren in het voortgezet onderwijs: Start, reacties en vervolg [School performance indicators in secondary education]. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: Scholen op rapport; standaarden voor de publicatie van schoolprestaties* (pp. 21-36). Assen, the Netherlands: Koninklijke Van Gorcum BV.

Early, P. (1998). *School improvement after inspection: School and LEA Responses*. London: Sage.

Elte, R. (1988). *Van het toezicht op het onderwijs: Over de inrichting en het functioneren van de inspectie in het algemeen voortgezet onderwijs en het lager beroepsonderwijs* [About school inspections: Set-up and functioning of the inspectorate in general secondary education and lower vocational education]. De Lier, the Netherlands: Academisch Boeken Centrum.

Educational Council. (2001). *De markt meester? Een verkenning naar marktwerking in het onderwijs* [Master of market? An exploration of market forces in education]. Den Haag, the Netherlands: Onderwijsraad.

Fleurke, F., & Huizenga, F. D. (1988). Methodische aspecten van beleidstheorie [Methodical aspects of policy theory]. *Bestuurswetenschappen*, *42*, 467-486.

Geijsel, F. (2001). *Schools and innovations: Conditions fostering the implementation of educational innovations*. Nijmegen, the Netherlands: University Press.

Gray, C., & Gardner, J. (1999). The impact of school inspections. *Oxford Review of Education*, *25*, 455-469.

Gray, J., & Wilcox, B. (1995). *Good school, bad school: Evaluating performance and encouraging improvement*. Buckingham, UK: Open University Press.

Hargreaves, D. H. (1995). Inspection and school improvement. *Cambridge Journal of Education*, *25*, 117-126.

Hoogerwerf, A. (1984). Beleid berust op veronderstellingen: De beleidstheorie [Policy is based on assumptions: Policy theory]. *Acta Politica: Tijdschrift Voor Politocologie*, *4*, 493-531.

Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behaviour in organizations. *Journal of Applied Psychology*, *64*, 349-371.

Karstanje, P. N. (1996). *Beleidstheorie basisvorming: Een proeve van reconstructie* [Policy theory of the basic education: An exercise of reconstruction]. Amsterdam: SCO rapport 412.

Karsten, S., & Visscher, A. (2001). Ervaringen met het openbaar maken van schoolprestaties in Engeland en Frankrijk [Experiences with publishing school performance indicators in England and France]. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: Scholen op rapport; standaarden voor de publicatie van schoolprestaties* (pp. 36-54). Assen, the Netherlands: Koninklijke Van Gorcum BV.

Kerr, S. (1975). On the folly of rewarding A, while hoping for B. *The Academy of Management Journal*, *18*, 769-783.

Kluger, A. A., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*, 254-284.

Kogan, M., & Maden, M. (1999). An evaluation of evaluators: The OFSTED system of school inspection. In C. Cullingford (Ed.), *An inspector calls: OFSTED and its effect on school standards* (pp. 9-32). London: Kogan Page.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, *24*, 5-20.

Mason, R. O., & Mitroff, I. I. (1981). *Challenging strategic planning assumptions: Theory, cases and techniques*. New York: John Wiley.

Ministry of Education, Culture and Science. (1999). *Varieteit en waarborg; voorstellen voor de ontwikkeling van toezicht op onderwijs* [Variety and guarantee; proposals for the development of school inspections]. Retrieved from http://www.minocw.nl/toezicht/nota2/03.html

Perrin, B. (1999). Performance measurement: Does the reality match the rhetoric? A rejoinder to Bernstein and Winston. *American Journal of Evaluation*, *20*, 101-111.

Reezigt, G. J. (2001). *A framework for effective school improvement: Final report of the ESI project*. Groningen, the Netherlands: GION, Institute for Educational Research, University of Groningen.

Rosenthal, L. (2004). Do school inspections improve school quality? OFSTED inspections and school examination results in the UK. *Economics of Education Review*, *23*, 143-151.

Shaw, I., Newton, D. P., Aitkin, M., & Darnell, R. (2003). Do OFSTED inspections of secondary education make a difference to GCSE results? *British Educational Research Journal*, *29*, 63-77.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, *18*, 277-310.

Toulmin, S. (1958). *The uses of argument*. Cambridge, UK: Cambridge University Press.

Toulmin, S. (1964). *The uses of argument* (Rev. ed). Cambridge, UK: Cambridge University Press.

van Thiel, S., & Leeuw, F. L. (2003). De prestatieparadox in de publieke sector [The performance paradox in the public sector]. *Beleidswetenschap*, *2*, 123-143.

Visscher, A. J. (2002). A framework for studying school performance feedback systems. In A. J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 41-75). Lisse, the Netherlands: Swets & Zeitlinger.

Walraven, G. (1991). *De kwaliteit van de beleidstheorie; visies uit de beleidspraktijk* [The quality of the policy theory: Visions from policy practice]. Enschede, the Netherlands: University of Twente, Faculteit der Bestuurskunde.

Waterreus, I. (2003). *Lessons in teacher pay: Studies on incentives and the labor market for teachers*. Amsterdam: ICO.

Weiss, C. H. (1997). Program theory-based evaluation: Past, present, and future. In D. J. Rog & D. Fournier (Eds.), *Progress and future directions in evaluation: Perspectives on theory, practice, and methods* (New Directions for Evaluation, *76*, 41-55). San Francisco: Jossey-Bass.

Wiebes, F. (Ed.). (1998). *Van buiten naar binnen: Waardevolle ideeën van buitenstaanders* [From outside to inside: Valuable ideas from outsiders]. Den Haag, the Netherlands: Delwel Uitgeverij BV.