

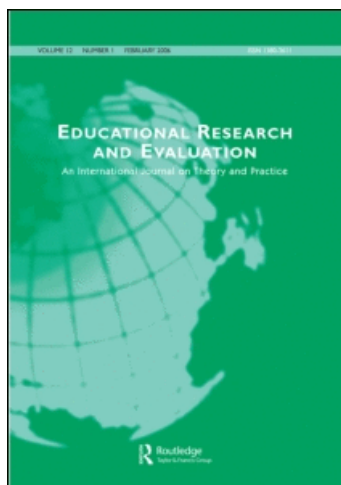
This article was downloaded by: [Universiteit Twente]

On: 9 December 2008

Access details: Access Details: [subscription number 789271096]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Educational Research and Evaluation

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t714592776>

School performance feedback systems in the USA and in The Netherlands: a comparison

Kim Schildkamp^a; Charles Teddlie^b

^a Department of Curriculum Design & Educational Innovation, Faculty of Behavioral Science, University of Twente, Enschede, The Netherlands ^b Department of Educational Leadership, College of Education, Louisiana State University, Baton Rouge, LA, USA

Online Publication Date: 01 June 2008

To cite this Article Schildkamp, Kim and Teddlie, Charles(2008)'School performance feedback systems in the USA and in The Netherlands: a comparison',Educational Research and Evaluation,14:3,255 — 282

To link to this Article: DOI: 10.1080/13803610802048874

URL: <http://dx.doi.org/10.1080/13803610802048874>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

School performance feedback systems in the USA and in The Netherlands: a comparison

Kim Schildkamp^{a*} and Charles Teddlie^b

^a*Department of Curriculum Design & Educational Innovation, Faculty of Behavioral Science, University of Twente, Enschede, The Netherlands;* ^b*Department of Educational Leadership, College of Education, Louisiana State University, Baton Rouge, LA, USA*

(Received 5 June 2007; final version received 25 October 2007)

Schools around the world are using instruments for performance feedback, but there is no scientific evidence that they have positive effects on education. This paper compares a School Performance Feedback System (SPFS) used in the USA as an accountability instrument to an SPFS used in The Netherlands. The study employs a unique database: one in which 2 separate countries with 2 distinct performance systems are compared using the same instruments. The use and effects of both SPFSs are compared to acquire more knowledge about the utilization and effects of SPFSs in an international context. Also, the variables which influence SPFSs are presented and then utilized to predict the use of the 2 SPFSs in their 2 separate contexts.

Keywords: School Performance Feedback Systems; School Improvement; accountability; evaluation

Introduction: general considerations regarding SPFSs

School Performance Feedback Systems (SPFSs) have become widespread in education all over the world (Coe & Visscher, 2002b). Research results show that, although schools around the world use SPFSs, performance feedback can both improve and harm performance (Coe, 2002). As stated by Coe and Visscher (2002a, p. 245): “enormous resources are invested to accomplish high quality SPFSs. However, the extent of any systematic check on how schools deal with the results of all that developmental work appears to be negligible.” Moreover, there is little knowledge of the factors that contribute to the use of SPFSs. A thorough evaluation of the use and effects of different SPFSs is urgently needed. Therefore, the central questions underlying this study are:

- (1) How and to what extent do schools use SPFSs?
- (2) What are the effects of the use of these SPFSs?
- (3) Which factors influence the use of SPFSs?

To gain insights into the use and effects of these SPFSs, an international comparative study into the use of a SPFS used in The Netherlands and a SPFS used in the USA was

*Corresponding author. Email: k.schildkamp@utwente.nl

conducted. As pointed out by, for example, Sammons (2006) and Creemers (2006) in the special issue of *Educational Research and Evaluation* on international studies on educational effectiveness, international studies are needed to understand the complex structure of education and to advance the development of theories about factors at and between different levels of the school system.

This article starts with describing a theoretical framework for studying SPFSs. Next, a Dutch SPFS, ZEBOS (the acronym stands for school self-evaluation in primary schools) is presented. The ZEBOS results reported in this article are based on a study conducted by Schildkamp (2007). The SPFS selected to evaluate in the USA (in Louisiana) is a SPFS used by the Louisiana Department of Education, the School Analysis Model (SAM). Thereafter, differences between SAM and ZEBOS are described, followed by the method used to answer the research questions. The article ends with the research results and conclusions and a discussion concerning the use of SPFSs in the USA and in The Netherlands.

Theoretical framework

According to Coe and Visscher (2002b), SPFSs are “information systems external to schools that provide them with confidential information on their performance and functioning as a basis for school self-evaluation” (p. xi). A theoretical framework for SPFSs developed by Visscher (2002) was applied in this study to systematically acquire detailed knowledge on the use and effects of ZEBOS and SAM.

Figure 1 displays this framework with all the variables studied and the assumed relationships between the groups of factors. As indicated in Figure 1, SPFS characteristics (A), the implementation process (B), and school organization characteristics (C) are all supposed to affect the use of the SPFSs (D). The use of the SPFSs is expected to generate intended and unintended effects (E).

The use of SPFSs

The first focus of this study is to investigate the use of SPFSs. The use of SPFSs, such as ZEBOS and SAM, may vary between schools and within schools. In some schools, the output generated by the SPFS may be studied only by the principal, or by individual teachers (Visscher, 2002).

Some schools may discuss the output widely, whereas other schools may not (Visscher, 2002). The output produced by the SPFS may provide school staff with new insights. Weiss (2001) concludes, from her study into the use of research results, that most results are not used directly and do not lead to changes in policy and practice, but they may challenge a lot of assumptions. Research results may undermine accepted myths and it may bring new ideas to the fore and change priorities. Weiss calls this “the enlightenment function of evaluations”.

The SPFS output may also highlight certain problems within the school. This may lead a school team to decide to attempt to devise solutions for the problems indicated. The final step is the implementation of these solutions (Visscher, 2002). The use of output generated by SPFSs may lead to certain (policy) measures at school level and at classroom level, taken by the whole school or individual teachers, with the ultimate goal of school improvement (Coe & Visscher, 2002a).

Following Weiss (1998a, 1998b), a distinction is made between instrumental use and conceptual use of the ZEBOS output. Rossi, Freeman, and Lipsey (1999) define instrumental

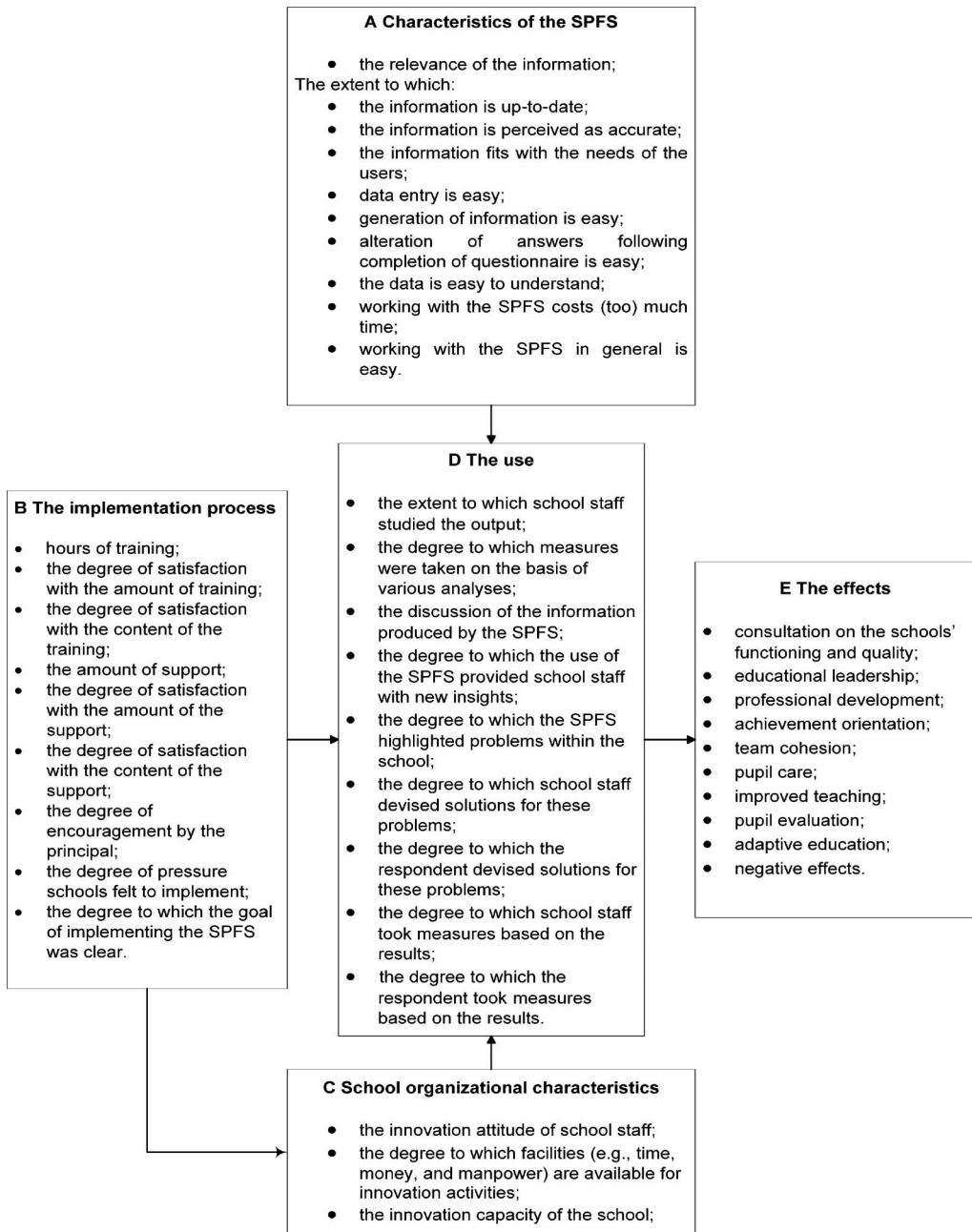


Figure 1. Factors expected to influence the use of SPFSs and its effects.

use as the direct use of evaluation findings: The results are analyzed and decisions and actions are based on the results. Conceptual use refers to the indirect use of evaluation findings. The performance feedback then influences thinking about issues in a general way and as such may have an (indirect) impact on the respondent's actions. This distinction between instrumental use and conceptual use will be utilized to study how schools use ZEB0.

The effects of SPFSs

To determine possible effects of ZEB0 and SAM, several variables were studied in depth. These variables were selected on the basis of the findings of school effectiveness and school improvement research. School effectiveness research is aimed at identifying those variables which are positively associated with pupil achievement. The aim of SPFSs such as ZEB0 and SAM is to improve the quality of the school, eventually in terms of pupil achievement. Therefore, the effects of the use of both SPFSs on those variables which seem to be positively associated with pupil achievement were included in this research. The following variables were selected, the effect on:

- the amount of consultation on the schools' functioning and quality. Without consultation and communication, it is very hard, if not impossible, to induce any change in an organization (Scheerens & Bosker, 1997);
- educational leadership. In several studies this characteristic is positively associated with pupil achievement (Leithwood, Jantzi, & Steinbach, 1999; Mortimore, 1998; Reynolds, Hopkins, Potter, & Chapman, 2002; Scheerens, 1990, 1991; Scheerens & Bosker, 1997; Witziers, Bosker, & Krüger, 2003);
- professional development of school staff. Kyriakides and Campbell (2004) state that school self-evaluation may stimulate professional development since it includes the systematic provision of feedback to staff and it may illuminate individual needs for professional development within the context of the school;
- the achievement orientation of school staff. A school staff with a high pupil achievement orientation is an important prerequisite for improved school performance (Reynolds et al., 2002; Scheerens & Bosker, 1997; Visscher, 2002);
- team cohesion. The degree of team cohesion is a school characteristic also consistently associated with high pupil achievement (Reynolds et al., 2002; Scheerens, 1990; Scheerens & Bosker, 1997);
- pupil care. Pupil care refers to the practice of attending to the personal and social well-being of children under the care of a teacher. Scheerens (1992) found, in his review of school effectiveness research, that intensive remediation, which may be considered an aspect of pupil care, is among the factors that best account for variation in the rate of learning of pupils;
- teaching, which is another requirement for improved school performance (Visscher, 2002). The (quality of the) functioning of teachers is very important in improving pupil achievement (Hoeben, 1995; Mortimore, 1998);
- evaluation. Frequent evaluation is also mentioned in school effectiveness research as a characteristic that is consistently positively associated with pupil achievement (Scheerens, 1990, 1991; Scheerens & Bosker, 1997);
- adaptive education. According to Houtveen, Booij, De Jong, and Van de Grift (1996, 1999), adaptive education is associated with higher pupil achievement. Adaptive education is education in which teachers, within a given context, direct their teaching to differences between pupils (Bolland, 1996, in Van den Berg & Vandenbergh, 1999).

Finally, the use of SPFSs may lead to certain negative effects. The (administrative) workload of teachers and principals may increase as a result of using a self-evaluation instrument (Van Petegem, 1998; Visscher, 2002), participants may feel threatened by the evaluation, and evaluations may evoke defensiveness (Clift, Nuttall, & McCormick, 1987),

and the evaluation may have a de-motivating impact on teachers, especially in poorly performing schools (Van Petegem, Vanhoof, Daems, & Mahieu 2005).

Factors influencing the use of SPFSs

In the research framework for this study, three groups of factors are assumed to influence the use of a SPFS: the characteristics of the SPFS, the features of the implementation process, and characteristics of the school organization in which the SPFS is implemented. These three groups of factors are discussed below.

Characteristics of the SPFS (A)

SPFSs may differ in the degree to which the output they provide is perceived as relevant by their users (Visscher, 2002). Teddlie, Kochan, and Taylor (2002) conclude, from their research into the ABC+ model for school diagnosis, feedback, and improvement, that the credibility and accessibility of the feedback for the users is very important.

The output produced by SPFSs may also differ in the degree to which the output is up-to-date (Visscher, 2002). Kimball (2002) calls this timeliness. After the gathering of the data, the feedback must be shared as soon as possible.

Furthermore, the degree to which the output is perceived as accurate and fits with the needs of the users (Visscher, 2002) may play a role in the use of the SPFS. Fullan (1991) states that the people involved must feel the need for the innovation. School staff must trust the output.

The user-friendliness of the system (Visscher, 2002) is also assumed to be an important factor in the use of SPFSs. It should not be too difficult to use the SPFS successfully; data entry, altering input, generating results, and interpretation of results should not be too complex. The synopses and statistics should not be too difficult to interpret, as the analysis and interpretation must be conducted by schools themselves (Coe & Visscher, 2002a).

Finally, the extent to which using SPFSs takes time and effort may differ between schools (Visscher, 2002). If the use of the SPFS is not perceived as difficult, schools are more likely to use the SPFS.

Implementation process features (B)

User training and support are assumed to play an important role in the successful implementation of a SPFS (Visscher, 2002). Users of SPFSs need certain skills and knowledge to successfully implement it, such as knowledge of how to interpret the statistics generated by the SPFS. Schools also may need support in the use of a SPFS, for example, internal support (e.g., from an employee that deals with ICT in the school) and external support (e.g., from the school counselling service). What is important is whether or not schools are satisfied with the amount of training and support they received (Visscher, 2002). If schools receive (in their opinion) enough training and support, they are more likely to use the SPFS intensively.

Moreover, if an implementation is actively encouraged and supported by the principal, the implementation is more likely to be successful. Fullan (1991) states, in this context, that the principal may shape the organizational conditions necessary for success, such as the development of shared goals, collaborative work structures, and procedures for monitoring results.

According to Visscher (2002), several authors believe that a combination of a pressure and support approach for implementing a SPFS will have the highest probability of success. Schools are more likely to improve their performances through the pressure of clear targets combined with external control. Fullan (1991) also states that both pressure and support are necessary for success. According to Gray (2002), once school staff feel some kind of external pressure, they will become more motivated to use the feedback produced by the SPFS.

Furthermore, Kyriakides and Campbell (2004) state similarly that it is important to establish clarity and consensus about the aims of the self-evaluation, starting with a clear understanding of the aims and of how the school self-evaluation will be conducted.

School organizational characteristics (C)

The degree to which schools and their staff possess the required attitudes, skills, and capacities for the innovation are considered important for schools using a SPFS (Visscher, 2002). Several authors (Rowe, Turner, & Lane, 2002; Teddlie et al. 2002; Tymms & Albone, 2002) stress the importance of a positive staff attitude towards the innovation. The attitude towards the innovation largely depends on the benefits and costs of the innovation as perceived by the school staff involved (Visscher, 2002).

Furthermore, school self-evaluation requires that schools devote a substantial amount of time, energy, and resources to it (Davies & Rudd, 2001; Kimball, 2002; Teddlie et al., 2002; Visscher, 2002). Schools are often very busy with routine activities that take up the most of their time. However, if schools have certain earmarked facilities (e.g., time, money, manpower) at their disposal for implementing innovations, this will probably lead to a more intensive use of the SPFS.

Another school organizational characteristic that may influence the degree of SPFS use concerns the innovation capacity of a school (Visscher, 2002). Geijsel, Van Den Berg, and Slegers (1996, 1999) conducted several studies into the implementation of innovations and the innovation capacity of schools. They define the innovation capacity as the capacity of schools to implement innovations in a successful manner.

The use of a SPFS in The Netherlands

In 1998, the Dutch “Quality Law” became operative. For schools and governing bodies, the most important stipulations relate to extending the competencies of the Inspectorate and to the so-called “principle of proportionality”. The latter means that the supervision of schools starts from the results of school self-evaluations (Inspectie van het Onderwijs, 2002; Ministerie van Onderwijs, Cultuur & Wetenschappen, 2000–2002; Renkema, 2002). Schools can use a SPFS to conduct a self-evaluation.

ZEB0 is one of those SPFSs the Dutch can use (the acronym stands for school self-evaluation in primary education in Dutch). This SPFS was chosen for two reasons. Firstly, it is a new system and it needed to be evaluated. Secondly, most available SPFSs in The Netherlands have serious technical weaknesses (reliability and validity issues), and the development of most instruments is not based on educational research. ZEB0 has been developed in a response to such a situation.

The development of ZEB0 took 5 years. Thirteen process variables that are often mentioned in school effectiveness research were selected for the development of ZEB0 (Scheerens & Bosker, 1997). The factors and components found on the basis of school effectiveness research were presented to primary principals and teachers. Based on their

opinions and on the conclusions of the committee for primary education evaluation in 1994, an initial selection of process variables was made. Table 1 displays an overview of the selected process variables and the level of measurement (Hendriks, Doolaard, & Bosker, 2002).

After two pilots (in 1997 and 1998), a final field test took place in 1999 in a representative sample of 123 schools in The Netherlands. In 2002, the final market version of ZEBO was released in a computerized form. This format allows schools to use ZEBO whenever they need the information, and they can obtain feedback immediately (Hendriks et al., 2002).

The process variables, displayed in Table 1 are measured by means of questionnaires for school management, teachers, and Grade 3–8 pupils. After completing the questionnaires in the schools, schools can generate two kinds of feedback (Hendriks et al., 2002):

- *A school report*: One can download graphic and written representations of the results of the school under study in comparison with schools from a national sample on each scale in the school report. Furthermore, the scores of the teachers are compared with the school management scores.
- *A classroom report*: This report is based on information from the pupil and teacher questionnaires. The results from the students of the school in a certain grade are compared to the results of students in the national sample from that same grade. The responses of the students are also compared with the responses of the teachers.

Table 1. Selected process variables and level of measurement (source: Hendriks, Doolaard, & Bosker, 2002, p. 124).

Process variable	Information collected at the level of		
	school management	teachers	pupils
School level			
- Achievement orientation/high expectations	X	X	X
- Educational leadership	X	X	
- Staff professionalization	X	X	
- Pupil care; measures that enable adaptive education	X	X	
Consensus and cohesion among staff:			
- frequency and content of formal staff meeting with school management	X	X	
- frequency and content of informal meeting among teachers (co-operation)	X	X	
School climate:			
- relationships between staff	X	X	
- relationship: the role of school management	X	X	
- workload	X	X	
Classroom level			
- Structured instruction		X	X
- Adaptive instruction			X
- Time on task			X
- Monitoring of pupils' progress		X	
- Pupil care: special care for high and low achievers		X	
- Classroom climate		X	X
- Relationships between pupils			X
- Teacher support and relationship between teacher and pupil			X

Schools have options about what kind of output they want to generate. With the feedback from ZEBO, schools may judge if and to what extent quality improvement is required and which activities are needed to improve their quality (Hendriks, 2001; Hendriks & Bosker, 2003). ZEBO is thus used for formative purposes by schools.

The use of a SPFS in the USA

Educational reform and accountability became a national priority in the USA with the publication of the Coleman report in 1966 (Coleman et al., 1966). Attention for educational reform increased with the publication of *A Nation at Risk* in 1983 (National Commission on Excellence in Education, 1983), and accountability became even more of a priority (e.g., Teddlie et al., 2002). Attention to accountability increased even more with the passing of the No Child Left Behind (NCLB) Act of 2001. The act requires states to implement state wide accountability systems.

A SPFS used in Louisiana for accountability purposes is the School Analysis Model (SAM). External assessors from the Louisiana Department of Education (a District Assistance Team) use SAM to conduct an on-site school evaluation. SAM incorporates school process data into a school accountability system. SAM is a model for gathering process data at the classroom, grade/department, and school levels and for interpreting and applying those data to the development of context-specific, school improvement planning. An ABC+ matrix was developed crossing four types of educational process variables by four sources of data (parent, student, classroom/teacher, school/principal).

The matrix provides for 16 data cells that can generate quantitative, qualitative, or mixed data. Definitions of the four types of educational process variables are:

- (1) attitudinal variables: emotions or feelings of individuals associated with a school;
- (2) behavioral variables: overt actions of individuals associated with a school;
- (3) cognitive variables: the level of cognitive functioning or understanding that individuals exhibit;
- (4) context variables: important school context variables including socioeconomic status of student body, community type, grade phase of schooling, and governance structure (Teddlie et al., 2002).

In SAM, external assessors conduct site visits to schools that fail to meet their growth targets: schools in School Improvement. A four- to five-member team of external assessors conduct a 2–3 day school site visit, using SAM protocols. The information collected is used to inform the design and delivery of follow-up technical assistance to the school and as a resource for school staff in their on-going school improvement planning (Teddlie et al., 2002).

These visits include a faculty needs assessment to gather information on the opinion and perceptions of the school; archival data; interviews with stakeholders (e.g., the principal, a school counselor, and teachers); stakeholder questionnaires (e.g., instructional staff, administrative staff, students, and parents); teacher, parent, and student focus groups; a contextual observation checklist; classroom observations; and an exit summary form (e.g., a preliminary report from the information gathered) (Louisiana Department of Education, 2002). Based on this, each school receives its top five strengths and weaknesses. This information should be used to develop a school's improvement plan (Louisiana Department of Education, 2002).

Differences and commonalities between SAM and ZEBO

SAM and ZEBO differ considerably but also have important commonalities. Table 2 shows the most important differences and commonalities.

Table 2. Differences and commonalities between SAM and ZEBO.

SAM	ZEBO
<i>The design and characteristics</i>	<i>The design and characteristics</i>
<ul style="list-style-type: none"> • Designed as a Professional Monitoring System • School effectiveness researchers, policy-makers at state level, and school staff were involved in the design of SAM • Offers a broad array of indicators. Schools can select the indicators relevant to their setting • Provides a school with information on the school process, the context, the input, and the output • Provides a school with qualitative and quantitative data • Both external and internal • Accountability driven • School improvement is the goal 	<ul style="list-style-type: none"> • Designed as a Professional Monitoring System • School effectiveness researchers and school staff were involved in the design of ZEBO • Offers a broad array of indicators. Schools can select the indicators relevant to their setting • Provides a school with information on the school process, the input, and the output • Provides a school with quantitative data • Internal driven • Improvement driven • School improvement is the goal
<i>The implementation process</i>	
<ul style="list-style-type: none"> • Great level of support in diagnosing and writing and implementing a school improvement plan. • The LDE is in-servicing all Louisiana public school staff on how to interpret data for school improvement. Hands-on staff development workshops have been conducted with all school principals and with school improvement teams from schools and districts state wide. 	<ul style="list-style-type: none"> • Low levels of support. It is the intention that schools diagnose and write and implement school improvement plans themselves. If necessary, schools can enlist the help of the school advisory service. • No in-services
<i>The school organization</i>	<i>The school organization</i>
<ul style="list-style-type: none"> • Schools with low performance levels 	<ul style="list-style-type: none"> • Schools with varying performance levels
<i>The use</i>	<i>The use</i>
<ul style="list-style-type: none"> • Used as an Official Accountability System • Members of the District Assistance Team (DAT) and school staff collect the data • DAT members analyze the data • External evaluators provide feedback on the data analyses • Provides schools with recommendations for improvement • Data collection involves teachers, school management, students, parents, and school counselors 	<ul style="list-style-type: none"> • Used as a Professional Monitoring System • School staff collects the data • School staff analyze the data • Schools staff have to provide the feedback • Provides schools only with the performance feedback • Data collection involves teachers, school management, and students

With regard to the communalities, the design of ZEBO and SAM and their characteristics are similar to a certain extent. Both ZEBO and SAM were developed as a Professional Monitoring System (PMS) to generate data for voluntary use by schools. Both instruments were developed on the basis of school effectiveness research and consultation with school staff and, in the case of SAM, consultation with policy-makers at state level. ZEBO and SAM both offer schools a broad array of indicators concerning school process, school input, and school output variables. SAM also provides schools with information on context variables, such as the state of the buildings. ZEBO and SAM both provide schools with quantitative data on their functioning, but SAM also provides schools with qualitative data, such as the results of the interviews. The last important similarity between ZEBO and SAM concerns the goal of both instruments. Schools should use the data gathered either through ZEBO or SAM for school improvement purposes.

With regard to the differences between ZEBO and SAM, one important difference between ZEBO and SAM is that ZEBO is an internal driven self-evaluation instrument and SAM is an accountability driven external evaluation instrument. Dutch schools, however, can and do use the self-evaluation as an accountability check for parents as well as for the school board. Louisiana schools are required to use the evaluation information to *improve* their functioning. One could argue that SAM is thus also improvement driven, but this improvement is again judged by the accountability system.

Although it can be argued that both SPFSs may be used for formative as well as summative purposes, widespread doubt exists that the summative and formative purposes can be achieved in a single set of evaluation arrangements (Kyriakides, 2005). As stated by Brown (1991), formative assessment requires school staff to be open and honest about success and failure, and summative assessment may lead to hide problems and teaching to the test to improve results. The results of this study may shed some light on the possible link between these two types of evaluation.

Furthermore, ZEBO and SAM differ extensively with regard to the implementation process. Schools receive a great level of support in using SAM. The external evaluators (District Assistance Team or DAT) collect most of the data (in cooperation with school staff), and the DAT helps schools in diagnosing and writing, and implementing a school improvement plan. Schools which are using ZEBO do not receive training or support in the use of the system. However, if necessary, schools can enlist the help of the school advisory service, an external agency in The Netherlands which can support schools in these kinds of issues.

Also, it can be argued that the pressure to improve will differ between schools that are using ZEBO and schools that are using SAM. Schools that are using SAM are more likely to feel greater pressure, because most of these schools failed to meet their growth targets. Nevertheless, some schools in Louisiana use SAM voluntarily and not because they are in School Improvement. Schools in The Netherlands are not obliged to use ZEBO.

With regard to SAM, external evaluators collect the data, analyze, and provide the school with feedback and recommendations for school improvement. School staff using ZEBO have to collect and analyze the data themselves. However, despite all these differences, both systems have one important aspect in common: Both systems are SPFSs.

Method and data sources

Characteristics of the samples from The Netherlands and the USA

The implementation of ZEBO was investigated in 70 primary schools in the Twente region of The Netherlands. All primary schools in this region were asked to participate on a voluntary basis. The sample schools have a smaller average school size ($F = 10.61$;

$p \leq 0.01$), but the sample is representative in terms of the composition of the student population (in terms of their socioeconomic status and ethnicity) ($F = 0.26$; $p = 0.61$). Schools started using ZEBO in 2002.

In 2004, 58 schools used ZEBO for the underlying study. Twelve schools did not use ZEBO for the following reasons:

- The school board or school chose to continue with another quality care instrument.
- Renovation of the school.
- Management problems or changes.

SAM has been used in Louisiana since 2000. The study described in this article involves the use of SAM in 2004, when 163 schools used SAM in Louisiana. Only 8 out of 36 districts were willing to cooperate in this study. They asked the 46 schools in their districts to participate, and 27 schools were willing to participate in this study. Twenty-three of these schools were in School Improvement (underperforming) and thus obliged to use SAM. Although only 27 schools participated in the study, the sample is representative of the composition of the student population, since there were no significant differences between the sample and the 163 schools in terms of several important characteristics:

- percentage of minority students ($F = 0.01$, $p = 0.94$);
- percentage of at-risk students ($F = 0.01$, $p = 0.92$);
- percentage of limited English speaking students ($F = 0.90$, $p = 0.34$);
- school size ($F = 0.14$, $p = 0.71$); and
- school type (primary school, high school, or combination school) ($F = 0.20$, $p = 0.65$).

The sample size from The Netherlands and the USA are both limited. Also, it must be taken into account that the Dutch sample is not representative for school size. Therefore, the results can not be generalized to all schools indiscriminately. However, valuable lessons can be learned. The sample size in the USA is smaller. Nevertheless, the sample is representative of other schools in Louisiana, and therefore the limited sample size was considered acceptable.

Also, the relatively small sample size is defensible, since international comparisons of SPFSs employing the same instruments (and researchers) are very rare. Yet, such international comparisons are necessary for the further development of the field.

Questionnaires used in the evaluation of the use of ZEBO and SAM

In order to study how schools use the SPFS output, what the effects of the SPFS are, and which factors influence the use of the SPFSs, a questionnaire was developed. The questionnaire was devised based on the framework underlying this study. All variables discussed above were included in the questionnaire. Two versions of the questionnaire were composed: one for principals and one for teachers. The two questionnaires are identical, with the exception of only five items which are appropriately phrased to address either teachers or principals.

The items in the questionnaire were designed to study the groups of factors in Visscher's (2002) framework:

- Characteristics of the SPFS: This scale is comprised of 9 items, assessing, for example, the perceived clarity and relevance of the output.

- Implementation process features: 7 items are included in this scale, rating such aspects as the clarity of the goal and the number of hours of training and support received.
- School organizational characteristics: 19 items were formulated for this scale, measuring, for example, the innovation attitude of staff and time and resources available for innovation activities.
- Conceptual SPFS use: This scale consists of 4 items evaluating, for example, new insights as a results of the use of a SPFS.
- Instrumental SPFS use: This scale consists of 5 items evaluating, for example, new measures taken based on the SPFS.
- The effects of SPFS use: This scale includes 9 items, appraising such effects of the SPFS use as, for example, those on pupil care and adaptive education.

For almost all questionnaire items, a statement format with a 4-point response scale, (ranging from 1 – *strongly agree* to 4 – *strongly disagree*) along with “I don’t know” and “does not apply” options, where appropriate, was provided. The direction of some items was reversed to prevent response bias. Table 3 presents more information on the factors in the questionnaire, including mean, standard deviation, and range.

We used the criteria developed by De Heus, Van der Leeden, & Gazendam (1995) to estimate the reliability of the ZEBO scales: The reliability of a scale is “good” if $\alpha \geq .80$ and “sufficient” if $.60 \leq \alpha < .80$. All but two of the ZEBO scales met the criteria (see Table 4). The exceptions were the scales “pressure and promoting factors” at teacher level and “innovation attitude” at the principal level. These scales were not used in the analyses. With regard to the questionnaire used in the Louisiana schools, all scales meet this criterion of reliability (see Table 4).

The 58 schools which had used ZEBO in 2004 were sent nine questionnaires each (one for the principal and one for each teacher from Grades 1 to 8) 3 months after they had used ZEBO. Fifty schools (86% response rate) returned some or all of the questionnaires, 3 to 6 months after they received the questionnaire. On average, six questionnaires were sent back from each school.

The Louisiana Department of Education (LDE) assisted in the data collection in the USA. Officials at the LDE sent the questionnaires out to the 36 school districts which had schools using SAM in 2004 (163 schools), with a request to each schools’ administrative staff to distribute the questionnaires in their schools (one questionnaire for the principal and one for each teacher from Grades 1 to 8). Eight districts were willing to cooperate and send the questionnaires (nine to each school) to 46 schools which had worked with SAM. Twenty-seven schools (59% response rate) send a total of 171 questionnaires back. On average, six questionnaires were sent back per school.

Table 3. Characteristics of the questionnaire scales.

	ZEBO <i>M</i> (items)	SAM <i>M</i> (items)	ZEBO <i>SD</i>	SAM <i>SD</i>	ZEBO range	SAM range
Characteristics of the SPFS	30.14 (9)	30.11 (9)	4.71	3.51	22.00	24.00
Implementation process features	25.82 (7)	21.54 (7)	4.72	2.93	17.00	19.00
School organizational features	64.18 (19)	59.35 (19)	6.65	4.80	32.00	23.00
Conceptual use of the results	6.80 (4)	9.51 (4)	2.51	2.59	9.00	11.00
Instrumental use of the results	10.41 (5)	14.53 (5)	3.71	2.94	13.00	14.00
Effects of the use of the SPFS	13.09 (9)	21.36 (9)	4.92	7.14	18.00	27.00

Table 4. Reliability of the ZEBO and SAM questionnaire scales at teacher and principal levels.

Reliability at	ZEBO: 2004 PL Cronbach's α (items)	ZEBO: 2004 TL Cronbach's α (items)	SAM: 2004 PL Cronbach's α (items)	SAM 2004 TL Cronbach's α (items)
<i>Scale</i>	<i>(N = 48)</i>	<i>(N = 236)</i>	<i>(N = 48)</i>	<i>(N = 236)</i>
Characteristics of the SPFS	0.73 (9)	0.80 (9)	0.73 (9)	0.77 (9)
Implementation process features: training and support	Too many missing cases	0.89 (4)	0.83 (4)	0.74 (4)
Implementation process features: pressure and promoting factors	0.65 (3)	0.59 (3)	0.65 (3)	0.61 (3)
School organizational features: innovation attitude	0.42 (7)	0.68 (7)	0.83 (7)	0.61 (7)
School organizational features: innovation capacity	0.78 (12)	0.84 (12)	0.80 (12)	0.81 (12)
Conceptual use of the results	0.72 (4)	0.75 (4)	0.72 (4)	0.74 (4)
Instrumental use of the results	0.66 (5)	0.75 (5)	0.66 (5)	0.75 (5)
Effects of the use of the SPFS	0.90 (9)	0.89 (9)	0.93 (9)	0.97 (9)

Note: PL: principal level, TL: teacher level.

The generalizability of the findings of this survey can be questioned, especially since in Louisiana, only 8 out of the 36 districts sent back questionnaires. However, compared to the usual response rate in educational studies and considering the fact that teachers and principals are oversubscribed to participate in research projects, the response rate was considered as adequate for ZEBO as well as for SAM.

Data analyses

With regard to the first research question, "How and to what extent do schools use SPFSs?", a distinction was made between the instrumental use and the conceptual use of the performance feedback from the SPFSs. Rossi et al. (1999) defined instrumental use as the direct use of evaluation findings: The results are analyzed and decisions and actions are based on the results. An example of an instrumental use item includes "the respondent took measures based on the performance feedback." Conceptual use, on the other hand, refers to the influence that performance feedback has on respondents' thinking about issues in a general way, which in the long run may have an impact on the respondent's actions. An example of a conceptual use item is "the SPFS results led to new insights." Answering the first research question led to a description and comparison of how and to what extent schools use ZEBO and SAM.

To answer the second research question, "What are the effects of the use of these SPFSs?", the data obtained by the questionnaires were used. The data for the intended and unintended effects were analyzed by means of frequencies. It was, for example, studied to what degree the use of ZEBO/SAM led to an effect on the functioning of the school leader

and to what degree the use of ZEBOSAM had negative effects. All the effects described in this article concern the perceived effects of the use of ZEBO and SAM.

To answer the third research question, "Which factors influence the use of SPFSs?", data from the questionnaires were first analyzed using correlations between the use of the SPFSs (ZEBO and SAM) (D) and the three groups of independent variables in Figure 1 (A, B, and C). To further investigate the extent to which the variance in ZEBO and SAM use could be explained by the independent variables, those variables which correlated significantly and strongly ($\alpha \leq 0.01$) with the use of ZEBO or SAM were entered into regression and multilevel analyses.

We recognize potential problems with statistical conclusion that this could cause (Shadish, Cook, & Campbell, 2002). Nevertheless, considering the number of variables in the theoretical framework that are all expected to influence the use of ZEBO and SAM, it was necessary to reduce the variables that were entered into the regression and multilevel analyses. Since it was impossible to come to a theoretical reduction because all variables seemed equally important, we decided to make a reduction based on empirical results from the correlational analyses.

Regression analyses were carried out on the data from the principals. Because the teacher data collected in this study have a nested structure (teachers are nested within schools), multilevel analysis was required for the teacher data (Schildkamp & Visscher, 2007).

Results

The use of ZEBO and SAM as perceived by its users

The first question asked "How and to what extent do schools use SPFSs?" Results reported in this section are descriptive in nature and relate to both the conceptual and instrumental use of the SPFSs as perceived by people using those systems. Thirty-four respondents (12%) from eight schools (14%) used ZEBO *conceptually* in 2004: The results influenced their thinking (i.e., the results increased the awareness on the weaker and stronger aspects of their functioning). The results, for example, made teachers more aware of their own functioning and the way that they are perceived by their pupils. The results pointed out specific problems, and the respondents came up with solutions for addressing these problems.

Forty-four respondents (26%) from eight schools (27%) indicated that they used the SAM results *conceptually*. New insights mentioned by the respondents included (1) the need to use more technology in the classroom and (2) the need for different and newer instructional methods. The problems found by SAM were related to these new insights: the need for technology use and different instructional methods in the classroom. Also, parental involvement and discipline needed improvement in several schools, according to the respondents.

Forty-three respondents (15%) from 11 schools (19%) used ZEBO *instrumentally*. The respondents indicated that they studied the ZEBO output and discussed the results. Measures that were taken to improve the quality of education included more frequent and open consultation about the quality of education, stimulating independent learning of pupils (e.g., by means of block teaching), more frequent evaluation of student achievement, more differentiation, implementing classroom consultation, using the results from ZEBO for the school plan and school prospectus, and more explanation in the classroom.

The SAM results were used *instrumentally* by 49 respondents (29%) from eight schools (27%). The SAM results were studied and discussed intensively in these schools.

Respondents mentioned that measures were taken to improve the quality of education based on these results. These measures included changes in the curriculum, such as the implementation of new reading and math programs and encouraging teachers who “try to go across the curriculum.” One teacher stated, for example, that she is trying to incorporate math issues in her social studies course. Furthermore, several schools took measures to try to involve the parents in the school, for example, by using newsletters and by organizing family nights at school.

Effects of the use of ZEBO and SAM as perceived by its users

The second question asked “What are the effects of the use of these SPFSs?”, and this question was also answered using descriptive statistics. The effects of the use of ZEBO and SAM that were mentioned most frequently included (1) an increase in the amount of consultation on the quality of education, (2) an increase in professional development activities, and (3) changes in the instructional strategies used by teachers. The exact results are presented in Tables 5 and 6.

Firstly, the largest improvement was found for the variable measuring “the amount of consultation on the quality of education” (1). This is very important, because without consultation and communication it is very hard, if not impossible, to induce any change in an organization (Scheerens & Bosker, 1997). Principals (38%) and teachers (20%) indicated that the use of ZEBO led to an increase in the amount of consultation on the quality of education to a moderate degree. Principals (35%) and teachers (30%) using SAM also indicated that the use of SAM also led to an increase in the amount of consultation to a moderate degree. An example mentioned here was that the mission statement of the school was discussed more often, both in The Netherlands and in Louisiana.

Secondly, respondents indicated that the use of ZEBO/SAM had an effect on the professional development of school staff (2). The use of the SPFSs stimulated professional development and illuminated individual needs for professional development within the context of the school. The professional development of school staff improved to a moderate degree in schools using ZEBO according to 25% of the principals and 9% of the teachers, and according to 26% of the teachers and 42% of the principals using SAM. Teachers and principals indicated, for example, that they took certain professional development courses, and teachers and principals started working with personal development plans.

Thirdly, the use of the SPFSs influenced teaching (3), which is another requirement for improved school performance (Visscher, 2002). The use of ZEBO led to changes in instruction methods to a moderate degree according to 5% of the teachers and 10% of the principals. Examples that were mentioned here included more classroom differentiation and stimulating independent learning. SAM users (35% and 62%, respectively) also mentioned this effect. Teachers started collaborating (more) and started to share instructional ideas to improve their teaching. Several teachers indicated that they started using more “hands-on” activities, implemented thematic lessons, incorporated technology use in their lessons, and started focusing more on critical thinking skills.

Principals and teachers using ZEBO did not report strong negative effects. Principals and teachers using SAM did not report any strong negative effects either. Only 5% of the teachers and 8% of the principals indicated that the use of SAM had led to negative effects (to a moderate degree). These negative effects were related to the “enormous amount of time” working with SAM had cost and the stress it had induced. Several teachers answered

Table 5. Perceived effects of the use of ZEB0 and SAM by teachers.

Effect on:	To a great degree % (n)		To a moderate degree % (n)		To a reasonable small degree % (n)		To a small degree/not % (n)		Missing % (n)	
	ZEB0	SAM	ZEB0	SAM	ZEB0	SAM	ZEB0	SAM	ZEB0	SAM
Amount of consultation	2 (4)	9 (13)	20 (47)	30 (43)	27 (63)	32 (47)	45 (106)	12 (18)	7 (16)	17 (24)
Professional development	1 (2)	10 (14)	9 (22)	26 (38)	14 (33)	27 (39)	58 (137)	26 (37)	18 (42)	12 (17)
Instructional strategies	0 (0)	8 (12)	5 (11)	35 (50)	12 (29)	28 (41)	69 (162)	13 (19)	14 (34)	16 (23)

Table 6. Perceived effects of the use of ZEBO and SAM by school leaders.

Effect on:	To a great degree		To a moderate degree		To a reasonable small degree		To a small degree/not		Missing	
	ZEBO	SAM	ZEBO	SAM	ZEBO	SAM	ZEBO	SAM	ZEBO	SAM
The amount of consultation	2 (1)	27 (7)	38 (18)	35 (9)	33 (16)	35 (9)	23 (11)	0 (0)	4 (2)	4 (1)
Professional development	0 (0)	23 (6)	25 (12)	42 (11)	31 (15)	23 (6)	40 (19)	12 (3)	4 (2)	0 (0)
Instructional strategies	0 (0)	8 (2)	10 (5)	62 (16)	29 (14)	19 (5)	54 (26)	8 (2)	6 (3)	4 (1)

to the question which measures they were taking to improve the functioning of the school that they focused their teaching mainly on the skills to be addressed in the state tests. This “teaching to the test” can be described as a negative consequence of the use of SAM.

Factors influencing the use of ZEBO and SAM: principals

The third question asked “Which factors influence the use of ZEBO and SAM?” Table 7 shows the significant regression weights in the regression analyses ($p < .05$) conducted to answer this question.

The following variables explain 31% of the variance in *conceptual* ZEBO use and 33% of the variance in *instrumental* ZEBO use, the degree to which:

- principals think the use of ZEBO will lead to quality improvement;
- the principal is not afraid of changes because of ZEBO;
- the information coincides with the needs of the users;
- the principal encourages the professional development of school staff.

Furthermore, Table 7 shows that four variables explain 59% of the variance in *conceptual* SAM use, and 60% of the variance in *instrumental* SAM use, the degree to which:

- extra resources were available for the use of SAM;
- the information coincides with the needs of the users;
- the school feels responsible for the education they provide;
- the principal feels that the school received enough training in the use of SAM.

Factors influencing the use of ZEBO and SAM: teachers' data

Multilevel analyses were conducted on the teacher data with the *conceptual* use of ZEBO or SAM as the dependent variable. Table 8 shows the results of these analyses. Two models are presented in Table 8: Model 0, which is the basic model without the independent variables, and Model 1. Model 1 is the same as Model 0, with the exception that the variables explaining the variance in the *conceptual* use of ZEBO or SAM have been added.

The model with the added ZEBO characteristics (A), implementation process (B), and school organizational (C) variables fit the data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2 = 50$, $df = 6$, $p < 0.01$) is significant. The model with the added implementation process (B) and school organizational (C) variables also fit the SAM data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2 = 42$, $df = 2$, $p < 0.01$) is significant. The following variables explain variance (27% between schools and 30% between teachers) in the *conceptual* use of ZEBO and variance (21% between schools and 36% between teachers) in the *conceptual* use of SAM, the degree to which:

- it is easy to input data;
- teachers are satisfied with the amount of training;
- the principal encouraged the use of the SPFS;
- teachers are not afraid of changes because of the use of the SPFS;

Table 7. Significant regression weights ($p < .05$) from regression analyses predicting conceptual use and instrumental use based on the principal data.

Independent variable	Coefficients	Conceptual ZEBU use ^a	Instrumental ZEBU use ^b	Conceptual SAM use ^c	Instrumental SAM use ^d
The information coincides with the needs	B		1.47	0.65	
	β		0.37	0.52	
	Sig.		0.03	0.00	
Enough training in the use of the SPFS	B				0.22
	β				0.56
	Sig.				0.00
SPFS use will lead to quality improvement	B	1.25			
	β	0.31			
	Sig.	0.03			
Not afraid of changes because of the SPFS	B	2.12			
	β	0.44			
	Sig.	0.00			
Extra resources available for use of the SPFS	B			-0.54	
	β			-0.55	
	Sig.			0.00	
Team is responsible for education	B				0.22
	β				0.38
	Sig.				0.03
Principal encourages professional development	B		2.50		
	β		0.35		
	Sig.		0.04		
	R ²	0.31 (0.00)	0.33 (0.00)	0.59 (0.00)	0.60 (0.00)

Note: ^a $n = 38$. ^b $n = 32$. ^c $n = 23$. ^d $n = 20$ (numbers differ from other tables due to the fact that not all principals answered all items needed for these analyses).

Table 8. Variables influencing the conceptual use of ZEB0 and SAM (teachers).

Fixed effects	Model 0 ZEB0 (N = 120)*		Model 1 ZEB0 (N 120)		Model 0 SAM (N = 145)		Model 1 SAM (N = 145)	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Intercept	-0.01	0.100	-0.05	0.09	0.08	0.15	0.05	0.10
Easy to input data			0.21	0.07				
Sufficient training			0.17	0.08			0.47	0.08
Encouragement use by Principal			0.19	0.08			0.30	0.08
Not afraid of changes			0.25	0.07				
Extra resources			0.16	0.07				
Teachers can influence measures			0.23	0.07				
Variance components								
Between schools	0.27	0.10	0.18	0.07	0.47	0.15	0.20	0.08
Between teachers	0.52	0.07	0.40	0.06	0.48	0.06	0.41	0.05
Percentage explained								
Between schools			27.40				21.19	
Between teachers			30.47				35.79	
Deviance	363		313		353		311	
Improvement in model fit (p)			0.00				0.00	

*(numbers differ from other tables due to the fact that not all teachers answered all items needed for these analyses).

- time and resources were available for the use of the SPFS;
- teachers feel that they can influence measures taken based on the SPFS results.

Multilevel analyses were also conducted on the teacher data with the *instrumental* use of ZEB0 and SAM as the dependent variable (see Table 9). Model 0 is the basic model, without the independent variables. Model 1 includes all significant independent variables for the instrumental use of ZEB0 and SAM, respectively.

The model with the added ZEB0 characteristics (A), implementation process (B), and school organizational (C) variables fit the data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2 = 40, df = 4, p < 0.01$) is significant. The model with the added implementation process (B) and school organizational (C) variables also fit the SAM data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2 = 58, df = 4, p < 0.01$) is significant. The following variables were found to explain variance in the instrumental use of ZEB0 and SAM (30% between schools and 30% between teachers in the use of ZEB0, and 28% between schools and 38% between teachers in the use of SAM), the degree to which:

- working with the SPFS costs a lot of time;
- hours of training were received in the implementation and use of the SPFS;
- teachers were satisfied with the support they received in the use of the SPFS;
- teachers felt forced to use the SPFS;

Table 9. Variables influencing the instrumental use of ZEBO and SAM.

	Model 0 ZEBO (<i>N</i> = 85)*		Model 1 ZEBO (<i>N</i> = 85)		Model 0 SAM (<i>N</i> = 85)		Model 1 SAM (<i>N</i> = 85)	
	Est.	<i>SE</i>	Est.	<i>SE</i>	Est.	<i>SE</i>	Est.	<i>SE</i>
Fixed effects								
Intercept	0.05	0.12	0.02	0.10	0.10	0.15	0.05	0.10
Time the SPFS costs			0.22	0.08				
Hours of training							0.25	0.07
Satisfied with support							0.16	0.07
Forced to use the SPFS							-0.24	0.07
Encouragement use by principal			0.23	0.09				
Teachers can influence measures			0.30	0.09			0.23	0.07
Principal encourages professional development			0.26	0.09				
Variance components								
Between schools	0.29	0.13	0.20	0.09	0.51	0.17	0.26	0.07
Between teachers	0.56	0.10	0.40	0.07	0.42	0.06	0.32	0.04
Percentage explained								
Between schools			30.24				28.46	
Between teachers			29.73				37.63	
Deviance	265		225		339		281	
Improvement in model fit (<i>p</i>)			0.00				0.00	

*(numbers differ from other tables due to the fact that not all teachers answered all items needed for these analyses).

- the use of the SPFS was encouraged by the principal;
- teachers feel they can influence measures taken based on the SPFS;
- the principal stimulates professional development.

Discussion and conclusions

Despite the fact that many SPFSs, such as ZEBO and SAM, have been developed and implemented, the number of evaluations conducted into the use and effects of these instruments are limited, and as a consequence a basis for improving (the use of) SPFSs is also missing (Coe & Visscher, 2002a). This study is unique in that it evaluates the use of two different SPFSs, used in two different countries in an international context.

By comparing the use of ZEBO and SAM, much can be learned about the use of SPFSs in schools, the effects of its use, and also with regard to the characteristics of schools which use SPFSs intensively as opposed to schools which do not, and as stated by Creemers (2006, p. 508) “we need more comparative international studies to gain a better understanding of what works in education and why.” By comparing these two SPFSs, we can learn what makes the use of these SPFSs more effective. Should schools be provided with more training, what is the role of the school leader in the use of these systems, are examples of questions that need to be answered.

In general, the research results found here are consistent with the theoretical framework underlying them. However, before we start generalizing these results, we

have to take into account a few contextual and methodological factors that are important to understanding this study.

First of all, in this study, teachers' and principals' self-perception is used to study the use and effects of SAM and ZEBO. These measurements may produce a slightly "colored" or biased picture of the actual use of ZEBO and SAM.

Secondly, we have to take into account that SAM is only used in Louisiana and that the use of ZEBO is studied in one region of The Netherlands. Consequently, the number of respondents participating in both studies was limited.

In addition, after the Louisiana schools completed SAM, Louisiana experienced two devastating hurricanes, Katrina and Rita. These hurricanes caused extensive damage, and all the schools participating in this study were directly or indirectly affected to some degree by these hurricanes. Several schools had, for example, to deal with an increase of student population due to closure of schools in the region destroyed by the hurricanes. This may have influenced the use of the SAM results. Schools may have had less time to use the SAM results, because they were, for example, busy with all kinds of organizational problems.

How can we compare the SPFS used in The Netherlands and the SPFS used in the USA? Despite the differences between both systems, there are also important commonalities, such as the fact that both systems can be described as SPFSs and both systems share the same overarching goal of school improvement. Data generated by ZEBO as well as SAM can be used by schools for feedback, diagnosis, and improvement. Most importantly, presumably all schools (whether underperforming or not), independently from their country, are interested in feedback on their performance in order to improve their performance. By comparing ZEBO and SAM, we can gain more insight into the use of SPFSs and come to conclusions on what works and what does not work for schools using a SPFS.

The use and effects of SPFSs

The results show that school staffs in Louisiana use the results slightly more than school staffs in The Netherlands for school improvement purposes (as perceived by the users). Nevertheless, it does not appear that schools in School Improvement (underperforming schools) in Louisiana use the results as much as one might have expected. Even though SAM has been used for several years now, it still appears to be difficult for school staff to use the SAM results to actually improve the quality of education. In line with other research (Coe, 2002; Weiss, 1998a, 1998b), we can conclude that it is and remains difficult to use evaluation results, whether these are self-evaluation results or external evaluation results.

With regard to the effects of SPFS use, not surprisingly considering the limited SPFS use, the effects of the use of both SPFSs are also limited so far. However, effects were found on important prerequisites for school improvement, such as improved communication, professional development, and teaching. SAM use had slightly larger effects than ZEBO use.

ZEBO is an internal improvement-driven instrument, whereas SAM is an external accountability-driven instrument. The question "can accountability and improvement be integrated in one system" is a complex question. Vanhoof and Van Petegem (2007) reflected on different types of integration of accountability and school improvement and concluded that integrating both types of evaluations is not without problems, due to the presence of different types of evaluatory functions, summative and formative functions.

Evaluation can have a formative function: using feedback for improvement purposes. Evaluation can also have a summative function: the evaluation of schools, for example, to study whether or not they are eligible for financing from the government. The purpose here is determining the results of a school. Vanhoof and Van Petegem (2007) further state that, when the functioning of the external evaluation is largely summative, using evaluation for accountability and improvement purposes is very problematic. Instead of focusing on improving the school, schools may be putting on a show to present themselves as positive as possible.

SAM can be described as both formative and summative in nature. Although the external audit is a summative report, schools can and have to use the feedback for improvement purposes. That this is not without problems, becomes clear from the SAM results. Several teachers, for example, started teaching to the test, to make sure that during the next evaluation their pupils would come out as positive as possible. The latter is mentioned by Brown (1991) as one of the dangers of the summative function of evaluations.

In the Dutch schools, external evaluation by the inspectorate is summative in nature. ZEBO is formative in nature. However, schools are responsible for their own quality assurance system and have to conduct some kind of school self-evaluation, used by the inspectorate as a starting point for inspection. Vanhoof and Van Petegem (2007) put forward the question whether or not self-evaluation can play a role in external evaluation at all, since there is the danger that school self-evaluation may be seen as an obligation and a task which is not connected to school improvement. The results from this study show that the ZEBO results are only used to some extent. This may very well have to do with the connection made by the inspectorate between external evaluation and self-evaluation.

In conclusion, the results of this study demonstrate that the relation between accountability and school improvement, internal and external evaluation, summative and formative evaluation is a very complex one. There is no easy answer as to how to integrate all these aspects. Both internal and external evaluations, summative and formative evaluations, and accountability and school improvement are important and necessary for an effective evaluation system. More research is needed on how to effectively integrate these different types of evaluations. Which type of approach (e.g., formative, summative, internal, external, accountability oriented, improvement oriented) is most effective in that it leads to more effective schools? Evaluating the use and effects of different evaluation approaches, focusing on using feedback for both school improvement and accountability purposes, can help strengthen and improve the SPFS framework underlying this study.

Kyriakides and Campbell (2004), for example, identify three possible ways to connect internal and external evaluations. In a parallel system, internal and external systems run side by side, each with their own criteria and protocols. In a sequential system, external bodies follow from a school's own self-evaluation, and in a cooperative system, external agencies cooperate with schools to develop a common approach. The Louisiana system can be described as a parallel system, whereas the Dutch system can be described as a sequential system. A question that needs more research is whether or not the use of a SPFS would increase in a more cooperative system. One could argue that important aspects, such as ownership, are more present in a cooperative system than in a sequential or parallel system.

The results of this study show that schools in The Netherlands and Louisiana did not make much use of the SPFSs, and the two systems did contribute significantly to school improvement processes within the school. However, given the (still growing) popularity of systems which provide school performance feedback to support schools in the process of

school self-evaluation and the fact that the results show that the use of these kinds of systems did have some (minor) effects, more research into the use and effects of these systems is justifiable. The factors included in this study may, to some extent, explain why schools did not use the performance feedback for improvement purposes.

Factors influencing the use of SPFSs

Education is a complex, contextually driven process. There is no “one-size-fits-all solution” for school improvement. Schools must solve their own problems and find their own unique way or path to improvement. SPFSs can help schools on this road to improvement. Although ZEBO is a self-evaluation instrument and SAM is an external evaluation instrument, similar factors prove to be important in the use of the evaluation results. These findings are important for those involved in the process of implementing school performance feedback systems.

Several variables were found to influence the use of the ZEBO results as well as the use of the SAM results, providing stronger empirical evidence for the importance of these variables. Knowledge on these variables can help us to develop more effective evaluation approaches. Firstly, schools that were more likely to use the feedback to make improvements were found to be those in which the SPFS information coincided with the needs of the users (*a SPFS characteristic*). Leithwood, Aitken, and Jantzi (2001) state in this light that it is important that an information need exists. Potential users should perceive a gap in their knowledge which the SPFS results can help fill. This finding is important for those involved in the process of implementing school performance feedback systems. Before implementing a system, one has to make sure that the feedback it provides coincides with the needs of the users.

Furthermore, *implementation process features* were found to influence the use of the SPFSs. Firstly, user training was found to play an important role. Saunders (2000) states, in this regard, that training and support which take into account individual school contexts are necessary for the data to be better understood and used. Scheerens, Glas, and Thomas (2003) state that, without external support, school self-evaluation is likely to fail. Schools often have problems in connecting “diagnosis” and “therapy”. This finding suggests that schools which are using SPFSs need training and support in the effective use of feedback in order for the feedback to lead to school improvement.

Secondly, the principal plays an important role in the use of SPFSs. It is important that the principal encourages the use of the SPFS and professional development of school staff in general. According to Kells (1995), unless the leader(s) of an organization are interested and willing to use the feedback, one should not proceed with the evaluation.

Organizational characteristics were also found to influence the use of both SPFSs. An increase in time and resources also led to an increase in ZEBO use. Davies and Rudd (2001) also found that the levels to which schools commit resources to self-evaluation, for example in the form of time and material support, plays an important role in the use of the self-evaluation results. However, the SAM results show that an increase in resources leads to a decrease of the use of the SAM results. It might be that these schools used those resources for other purposes than SAM. However, this finding may also be caused by measurement flaws. In general, it costs time and resources to use feedback effectively.

This study also demonstrates the importance of ownership in the use of evaluation results. Teachers who indicated that they felt that they were able to influence these used the SPFS results more. Several authors (Davies & Rudd, 2001; Kyriakides & Campbell, 2004) conclude that it is important to promote ownership of the (results of the) evaluation

among teachers. Teachers should have the opportunity to focus on aspects of the school that they identify as areas of improvement and take measures accordingly.

One can argue that developing ownership is easier in the case of ZEBO, since ZEBO is a self-evaluation instrument schools themselves selected to use. Most schools using SAM were more or less forced to use SAM. However, the study conducted by Schildkamp (2007) shows that in many schools, ZEBO was selected by the school board or school leader. Moreover, teachers indicated feeling pressured into using a self-evaluation instrument by the Dutch inspectorate. This makes the concept of ownership, for both internal and external evaluations, problematic, as was also stated by Kyriakides and Campbell (2004).

The question “how the use of feedback can be promoted in schools?” is partly answered by the results of this study. Several factors were found to influence the use of the SPFSs. However, part of the variance in the use of the SPFS results remains unexplained. This may partly be due to the limited SPFS use in most schools, leading to little variance in the use of the SPFSs. It may also be due to flaws of this study.

Another potential cause for the unexplained variance in SPFS use is that the framework developed by Visscher (2002) does not include all the variables relevant for studying the use of self-evaluation results. Research into some other independent variables may increase the explained variance in the use of school performance feedback systems. One can, for example, think of personality characteristics that may influence the willingness to use feedback such as the degree to which performance feedback is viewed by a principal or a teacher as either a threat or as an opportunity to improve his/her performance.

Since school self-evaluation and development of SPFSs receive much attention around the globe, the justifications for using these instruments are plausible. As thousands of schools are using SPFSs, further research into the use of these systems and into the conditions promoting effective SPFS is desirable.

Notes on contributors

Kim Schildkamp is an assistant professor at the department of Curriculum Design & Educational Innovation at the University of Twente. Her PhD thesis was on the use and effects of school self-evaluation. In 2005, she received a Fulbright scholarship to study the use of performance feedback in Louisiana. Currently, her research focus is on data-based decision-making for curriculum development.

Charles Teddlie is the Jo Ellen Levy Yates Professor (Emeritus) in the College of Education at Louisiana State University. He is the author of numerous books, chapters, and articles on research methods and school/teacher effectiveness including (with Abbas Tashakkori) *The Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Techniques in the Social and Behavioral Sciences* (Sage Publications, 2008).

Acknowledgements

With special thanks to the Fulbright Organization, and the Louisiana Department of Education.

References

- Brown, M. (1991). Problematic issues in national assessment. *Cambridge Journal of Education*, 21(2), 215–230.
- Clift, P.S., Nutall, D.L., & McCormick, R. (1987). *Studies in school self-evaluation*. Philadelphia: The Falmer Press.
- Coe, R. (2002). Evidence on the role and impact of performance feedback in schools. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 3–26). Lisse, The Netherlands: Swets & Zeitlinger.

- Coe, R., & Visscher, A.J. (2002a). Drawing up the balance sheet for school performance feedback systems. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 221–254). Lisse, The Netherlands: Swets & Zeitlinger.
- Coe, R., & Visscher, A.J. (2002b). Introduction. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. xi–xix). Lisse, The Netherlands: Swets & Zeitlinger.
- Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Creemers, B.P.M. (2006). The importance and perspectives of international studies in educational effectiveness. *Educational Research and Evaluation, 12*, 499–511.
- Davies, D., & Rudd, P. (2001). *Evaluating school self-evaluation*. Berkshire, UK: National Foundation for Educational Research.
- De Heus, P., Van der Leeden, R., & Gazendam, B. (1995). *Toegepaste data-analyse. Technieken voor niet-experimenteel onderzoek in de sociale wetenschappen* [Applied data-analysis. Methods for non-experimental research in social sciences]. 's-Gravenhage, The Netherlands: Elsevier Bedrijfsinformatie.
- Fullan, M.G. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Geijssel, F., Van Den Berg, R., & Slegers, P. (1996). Het innovatief vermogen van schoolleiders in het basisonderwijs: Een tweede vooronderzoek [The innovation capacity of school leaders in primary education: A second pilot study]. *Pedagogische Studiën, 73*, 42–55.
- Geijssel, F., Van Den Berg, R., & Slegers, P. (1999). The innovative capacity of schools in primary education: A qualitative study. *Qualitative Studies in Education, 12*(2), 175–191.
- Gray, J. (2002). Jolts and reactions: Two decades of feeding back information on schools' performance. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 143–162). Lisse, The Netherlands: Swets & Zeitlinger.
- Hendriks, M. (2001). Systemen voor kwaliteitszorg in het primair en secundair onderwijs [Systems for quality care in primary and secondary education]. In R.J. Bosker (Ed.), *Kwaliteitszorg. Onderwijskundig Lexicon III* (pp. 47–72). Alphen aan den Rijn, The Netherlands: Samson.
- Hendriks, M., Doolaard, S., & Bosker, R.J. (2002). Using school effectiveness as a knowledge base for self-evaluation in Dutch schools: The ZEBO-project. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 115–142). Lisse, The Netherlands: Swets & Zeitlinger.
- Hendriks, M.A., & Bosker, R.J. (2003). *ZEBO instrument voor zelfevaluatie in het basisonderwijs. Handleiding bij een geautomatiseerd hulpmiddel voor kwaliteitszorg in basisscholen* [ZEBO instrument for self-evaluation in primary schools. Manual for a computerized instrument for quality care in primary schools]. Enschede, The Netherlands: Twente University Press.
- Hoeben, W.Th.J.G. (1995). Integrale kwaliteitszorg [Integral quality care]. In J. Lowyck & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals* (pp. 287–314). Groningen, The Netherlands: Wolters-Noordhoff.
- Houtveen, A.A.M., Booij, N., De Jong, R., & Van de Grift, W.J.C.M. (1996). Adaptief onderwijs en leerlingresultaten [Adaptive education and student achievement]. *Pedagogische Studiën, 73*, 422–433.
- Houtveen, A.A.M., Booij, N., De Jong, R., & Van de Grift, W.J.C.M. (1999). Adaptive instruction and pupil achievement. *School effectiveness and school improvement, 10*, 172–192.
- Inspectie van het onderwijs. (2002). *Onderwijsverslag over het jaar 2001* [Education report from the year 2001]. Utrecht, The Netherlands: Author.
- Kells, H.R. (1995). Creating a culture of evaluation and self-regulation in higher education organizations. *Total Quality Management, 6*(5&6), 457–467.
- Kimball, S.M. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal of Personnel Evaluation in Education, 16*(4), 241–268.
- Kyriakides, L. (2005). Evaluating school policy on parents working with their children in class. *The Journal of Educational Research, 98*(5), 281–298.
- Kyriakides, L., & Campbell, R.J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation, 30*, 23–36.
- Leithwood, K., Aitken, R., & Jantzi, D. (2001). *Making schools smarter. A system for monitoring school and district progress*. Thousand Oaks, CA: Corwin Press.
- Leithwood, K., Jantzi, D., & Steinbach, R. (1999). *Changing leadership for changing times*. Philadelphia: Open University Press.

- Louisiana Department of Education. (2002). *The School Analysis Model (SAM 2000). User's guide for District Assistance Teams*. Baton Rouge, LA: Louisiana Department of Education.
- Ministerie van Onderwijs, Cultuur & Wetenschappen. (2000–2002). *Wet op het Onderwijstoezicht* [Education Supervision Act]. Den Haag, The Netherlands: SDU.
- Mortimore, P. (1998). *The road to improvement. Reflections on school effectiveness*. Lisse, The Netherlands: Swets & Zeitlinger.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for education reform*. Washington, DC: U.S. Government Printing Office.
- No Child Left Behind Act of 2001*, Pub. L. No. 107–110, 115 Stat. 1425 (2002).
- Renkema, G. (2002). Kwaliteitszorg, wat levert het op? [Quality Assurance – What are the gains?]. *Basisschoolmanagement*, 15(8), 14–20.
- Reynolds, D., Hopkins, D., Potter, D., & Chapman, C. (2002). *School improvement for schools facing challenging circumstances: A review of research and practice*. London: OFSTED.
- Rossi, P.H., Freeman, H.E., & Lipsey, M.W. (1999). *Evaluation. A systematic approach*. Thousands Oaks, CA: Sage.
- Rowe, K.J., Turner, R., & Lane, K. (2002). Performance feedback to schools of students' year 12 assessments: The VCE data project. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 163–190). Lisse, The Netherlands: Swets & Zeitlinger.
- Sammons, P. (2006). The contribution of international studies on educational effectiveness: Current and future directions. *Educational Research and Evaluation*, 12, 583–593.
- Saunders, L. (2000). Understanding schools' use of "value added" data: The psychology and sociology of numbers. *Research Papers in Education*, 15(3), 241–258.
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1, 61–80.
- Scheerens, J. (1991). Evaluation studies. Process indicators of school functioning: A selection based on the research literature on school effectiveness. *Studies in Educational Evaluation*, 17, 371–403.
- Scheerens, J. (1992). *Effective schooling, research, theory and practice*. London: Cassell.
- Scheerens, J., & Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science.
- Scheerens, J., Glas, C., & Thomas, S.M. (2003). *Educational evaluation, assesment, and monitoring*. Lisse, The Netherlands: Swets & Zeitlinger.
- Schildkamp, K. (2007). *The utilisation of a self-evaluation instrument for primary education*. Enschede, The Netherlands: Ipskamp.
- Schildkamp, K., & Visscher, A.J. (2007). *Factors influencing the utilization of a school self-evaluation instrument*. Manuscript submitted for publication.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for general causal inference*. Boston: Houghton Mifflin.
- Teddlie, C., Kochan, C., & Taylor, D. (2002). The ABC + model for school diagnosis, feedback, and improvement. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback*. (pp. 75–114). Lisse, The Netherlands: Swets & Zeitlinger.
- Tymms, P., & Albone, S. (2002). Performance indicators in primary schools. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 191–218). Lisse, The Netherlands: Swets & Zeitlinger.
- Van den Berg, R., & Vandenbergh, R. (1999). *Succesvol leiding geven aan onderwijsinnovaties. Investeren in mensen* [Successfully leading educational innovations: Investing in people]. Alphen aan den Rijn, The Netherlands: Samsom.
- Vanhoof, J., & Van Petegem, P. (2007). Matching internal and external evaluation in an era of accountability and school development: Lessons from a Flemish perspective. *Studies in Educational Evaluation*, 33, 101–119.
- Van Petegem, P. (1998). *Vormgeven aan schoolbeleid. Effectieve-scholenonderzoek als inspiratiebron voor de zelfevaluatie van scholen* [Forming school policy: Effective school research as an inspiration for school self-evaluation]. Leuven, Belgium: Acco.
- Van Petegem, P., Vanhoof, J., Daems, F., & Mahieu, P. (2005). Publishing information on individual schools. *Educational Research and Evaluation*, 11, 45–60.
- 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. 41–72). Lisse, The Netherlands: Swets & Zeitlinger.

- Weiss, C.H. (1998a). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Weiss, C.H. (1998b). Improving the use of evaluations: Whose job is it anyway? In A.J. Reynolds & H.J. Walberg (Eds.), *Advances in Educational Productivity* (Vol. 7, pp. 263–276). Greenwich/London: JAI Press.
- Weiss, C.H. (2001). *What kind of evidence in evidence-based policy?* Paper presented at the Third International, Inter-Disciplinary Evidence-Based Policies and Indicator Systems Conference, Durham, UK.
- Witziers, B., Bosker, R.J., & Kruger, M.L. (2003). Educational leadership and student achievement: The elusive search for an association. *Educational Administration Quarterly*, 39(3), 398–425.