

**THE UTILISATION OF A SELF-EVALUATION
INSTRUMENT FOR PRIMARY EDUCATION**

Kim Schildkamp

DOCTORAL COMMITTEE

Chairman: Prof. dr. H.W.A.M. Coonen ^a University of Twente

Promotor: Prof. dr. J. Scheerens ^a University of Twente

Assistant promotor: Dr. A.J. Visscher ^a University of Twente

Members: Prof. dr. J.J.H. van den Akker ^a University of Twente
Dr. R.H. Hofman ^a University of Groningen
Prof. dr. F.J.G. Janssens ^a University of Twente
Prof. dr. J.F.M. Letschert ^a University of Twente
Prof. dr. K. Sanders ^a University of Twente
Prof. dr. Prof. dr. P.J.C. Slegers ^a University of Amsterdam

The utilisation of a self-evaluation instrument for primary education
Thesis University of Twente, Enschede – With refs. – With Dutch summary.
ISBN 978-90-365-2466-7

Press: PrintPartners Ipskamp – Enschede

Cover design: Ilona Oosterik en Erik Post (Photos: Educatorium in Ootmarsum and
Drienermarke in Hengelo)

Lay-out: Sandra Schele

© Copyright, 2007, Kim Schildkamp.

e-mail: k.schildkamp@utwente.nl

The research was carried out in the context of the Interuniversity Center for Educational
Research (ICO) in the Netherlands.

*All rights reserved. No part of this book may be produced in any form: by print, photocopy,
microfilm, or any other means without written permission from the author.*

THE UTILISATION OF A SELF-EVALUATION INSTRUMENT FOR PRIMARY EDUCATION

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. W.H.M. Zijm,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 15 maart 2007 om 13.15 uur

door

Kim Schildkamp

geboren op 10 september 1979

te Hengelo (o)

Dit proefschrift is goedgekeurd door de promotor Prof. dr. J. Scheerens en de assistent-promotor dr. A.J. Visscher.

Voorwoord

Vaak kreeg ik tijdens mijn promotie traject de vraag “Is het niet saai om je vier jaar lang met hetzelfde onderzoek bezig te houden?”. Hier kon ik altijd maar één ding op antwoorden: “NEE”! Ik heb de afgelopen vier jaren niet alleen veel geleerd, maar ook veel plezier gehad in bijvoorbeeld het interviewen van leerkrachten en directeuren, het bezoeken van conferenties en zelfs het uitvoeren van allerlei statistische analyses.

In het voorwoord is het de gewoonte om mensen te bedanken die direct of indirect hebben geholpen bij het tot stand komen van dit proefschrift. Ik heb gezocht naar een originele manier om dit te doen, maar deze helaas niet kunnen vinden. Ik ga dus maar gewoon beginnen met het bedanken van deze mensen. Allereerst wil ik graag mijn begeleiders Jaap Scheerens en Adrie Visscher bedanken. Zonder hen had er geen proefschrift tot stand kunnen komen. Jaap, het was in de kerstvakantie nog even hard werken maar het is gelukt! Adrie, erg fijn dat je deur de afgelopen vier jaren altijd open stond!

Verder wil ik graag een aantal (in sommige gevallen oud-) collega's, Rolinda, Marloes, Karin, Gerdy, Martina, Maria, Rien, Lisenka, Carola, Karin, Leandra, Karlijn, Ida, Bob, Marinka, Melanie, Anouk, Ralf en Hans bedanken voor al hun hulp. Een aantal mensen wil ik nog even specifiek noemen. Maria, laat ik beginnen bij de basis: Als jij niet (mede) ZEBO ontwikkeld had, had ik dit onderzoek natuurlijk niet kunnen doen. Maar daarnaast heb jij de afgelopen jaren veel meer voor mij gedaan. Bij jou kon ik altijd terecht voor vragen, maar ook gewoon om even te kletsen. Rien, jij had en hebt altijd het antwoord op al mijn statistiek vragen en je hebt mij enorm geholpen bij mijn data verzameling. Hans, jij liet mij zien dat multilevel analyses ook leuk kunnen zijn. Lisenka, zonder jouw hulp had ik niet naar Amerika gekund. Carola, bedankt voor al je ondersteuning en we blijven fietsmaatjes!

Dan wil ik ook nog graag een aantal andere collega's bedanken. Allereerst de GW en ICO promovendi die ik in de afgelopen jaren heb leren kennen. Ik heb veel geleerd van onze werkgerelateerde bijeenkomsten, maar ook van de niet werkgerelateerde bijeenkomsten, zoals de aio-ententjes. Ook de chococlub, Cindy, Inge, Elvira, Carolien, Derk-Jan en aanhang, wil ik graag bedanken voor alle steun en gezelligheid. Ook mijn nieuwe C&O collega, Sandra Schele, wil ik bij deze graag bedanken voor het helpen bij het drukklaar maken van mijn proefschrift.

Tevens wil ik de mensen van Expertis en “mijn” scholen bedanken voor al hun hulp en medewerking bij dit onderzoek. Ondanks dat ik veel van “mijn” scholen vroeg (ZEBO gebruiken, vragenlijsten invullen, toetsgegevens opsturen, interviews) hebben jullie mij toch gegeven wat ik nodig had om dit onderzoek tot een goed eind te kunnen brengen. Ik heb veel geleerd van “mijn” scholen.

I also would like to thank several non-Dutch speaking colleagues: Sue Lasky, Charles Teddlie, Martha Hocutt, and Tasha Anthony. I learned so much from you. Also, Sheena Leeson, thank you for editing my thesis.

De belangrijkste mensen in mijn leven, mijn familie, wil ik ook graag bedanken in dit voorwoord. Een aantal familieleden wil ik graag specifiek noemen. Mam, zonder jou had ik dit nooit gekund! Pap, helaas mag je dit niet meer meemaken, maar ik draag dit proefschrift op aan jou. Mijn grote kleine broertje, Remi, ik ben trots op je! Gerrie, Gerrit, Gwen, Ria, Gerrit, Nico, Erik en oma: Bedankt voor het feit dat jullie altijd voor mij klaar staan. De leuke dingen die wij altijd samen doen kan ik aanbevelen als welkome afwisseling op het schrijven van een proefschrift!

Dan, last but definitely not least, wil ik mijn vrienden bedanken. Naast mijn familie zijn zij de belangrijkste personen in mijn leven en zonder hen was ik nooit zo ver gekomen. Ook hier wil ik graag een aantal mensen specifiek noemen. Harriët, het was de afgelopen jaren erg handig dat ik altijd bij je terecht kon voor wat “Engels advies”. Daarnaast wil ik je bedanken voor het feit dat ik al meer dan 20 jaar altijd bij je terecht kan. Marjolein, mijn “redding” toen het ging om het schrijven van de Nederlandse samenvatting en iemand die altijd bereid is om naar mijn verhalen te luisteren, waarvoor bedankt. Joke, in Nijmegen hebben we elkaar ontmoet. Sindsdien is er veel veranderd, maar ik ben erg blij dat we nog steeds goede vriendinnen zijn en bedankt voor je hulp bij het afronden van mijn proefschrift. Ilona en Arjan, jullie hebben de afgelopen jaren gezorgd voor veel gezelligheid, wat absoluut noodzakelijk is bij het schrijven van een proefschrift. Ook jullie wil ik bedanken voor jullie hulp bij het afronden van mijn proefschrift. Ik vind het geweldig dat jullie mijn paranimfen zijn! Mijn twee oud-kamergenootjes, Elvira en Chantal, wil ik bedanken voor alle hulp en gezelligheid de afgelopen jaren. Cindy, jouw kritische blik op mijn stukken kwam meer dan eens goed van pas. Verder was het ontzettend leuk om samen met jou te koken, te eten, film te kijken en te rocken. Hopelijk blijven we dit in de toekomst nog veel samen doen!

Kim

Voor papa

Table of Contents

1. Self-evaluation in Dutch Primary Schools: Introduction and Research Questions	1
1.1 Introduction	1
1.2 School Quality and School Quality Care	1
1.3 School Self-Evaluation	3
1.4 Self-Evaluation in Dutch Primary Schools	6
1.5 ZEBO	7
1.6 Research Questions	9
1.7 Overview of this Study	10
2. Theoretical Framework	11
2.1 School Performance Feedback Systems	11
2.1.1 A Theoretical Framework for Studying School Performance Feedback Systems	11
2.2 The Use of School Performance Feedback Systems	12
2.3 Effects of School Performance Feedback Systems	13
2.4 Factors Influencing the Use of School Performance Feedback Systems	16
3. Method	21
3.1 Research Design	21
3.1.1 Sample	21
3.2 Data Collection and Instruments	23
3.2.1 ZEBO	23
3.2.2 The Cito Pupil Monitoring System	25
3.2.3 The Pupil Form	26
3.2.4 Evaluation of ZEBO Questionnaire	27
3.2.5 Interviews	30
3.2.6 Mixed Methods	33
3.3 Data-Analyses	34

4. The Use of ZEBO	41
4.1 Introduction	41
4.2 ZEBO Use: The Results from the Evaluation of ZEBO Questionnaire	42
4.2.1 Conceptual Use of ZEBO output in 2003	42
4.2.2 Conceptual Use of ZEBO output in 2004 and 2006	43
4.2.3 Instrumental Use of ZEBO output in 2003	44
4.2.4 Instrumental Use of ZEBO output in 2004 and 2006	46
4.3 ZEBO Use: The Results from the Interviews in 2003 and 2005	49
4.4 A Comparison of the First, Second and Third Evaluations of ZEBO Use	52
4.5 Conclusions	54
5. The Effects of ZEBO use	57
5.1 Introduction	57
5.2 Effect of ZEBO Use on Pupil Achievement: Multilevel Analyses with Repeated Measures	58
5.2.1 The Effect of ZEBO Use on Pupil Achievement Growth for Cohort 1	61
5.2.2 The Effect of ZEBO Use on Pupil Achievement Growth for Cohort 2	63
5.3 Effects of ZEBO Use on Pupil Achievement: Conclusion	66
5.4 Perceived Effects of Using ZEBO: Results of the Evaluation of ZEBO Questionnaire	66
5.4.1 Results of the Evaluation of ZEBO Questionnaires in 2003	67
5.4.2 Results of the Evaluation of ZEBO Questionnaires in 2004 and 2006	69
5.5 A Comparison of the Perceived Effects of ZEBO Use on the Process Indicators	71
5.6 Perceived Effects of ZEBO use on Process Indicators: Conclusion	73

6. Factors Influencing the Use of ZEBO	75
6.1 Introduction	75
6.2 Factors Influencing the Use of ZEBO: Results of the Principals	76
6.2.1 Analysis of Correlations on the Data from the Principals	76
6.2.2 Multiple Regression Analyses Based on the Data from the Principals	78
6.3 Factors Influencing the Use of ZEBO: Results of the Teachers	80
6.3.1 Analysis of Correlations on the Data from the Teachers	80
6.3.2 Multilevel Analyses Based on the Data from the Teachers	83
6.4 The Results of the Interviews in 2003, 2004, and 2006	86
6.5 Conclusions	89
7. Conclusion and Discussion	97
7.1 Background, Research Questions and Method	97
7.2 Conclusions	98
7.2.1 How and To What Extent Do Schools Use ZEBO?	98
7.2.2 What Are the Intended and Unintended Effects of the Use of ZEBO?	99
7.2.3 Which Factors Influence the Use of ZEBO?	100
7.3 Reflections on the Research Design	103
7.4 Implications of the Study	103
7.4.1 Implications for Schools	103
7.4.2 Implications for School Self-Evaluation Designers	105
7.4.4 Theoretical Implications	105
7.5 Suggestions for Further Research	108
Summary	111
Samenvatting (Dutch Summary)	121
References	131
Appendices	141

Definitions

- § ZEB0 administration: the use of the ZEB0 instrument itself in schools. Schools used ZEB0 in 2003, 2004, and 2006.
- § ZEB0 output: this term refers to the output generated by the ZEB0 instrument. ZEB0 may produce two types of output: firstly, a school report which contains both graphic and written representations of the results for each scale for the school, in comparison with a national sample. The results of the teachers are also compared with the results of the principal. Secondly, a classroom report which contains a graphic representation and a textual explanation of the results of the pupils and teachers. The results of pupils of a certain grade are compared to the results of pupils in the national sample from that same grade. The results of the pupils are also compared with the results of the teachers.
- § The questionnaire: refers to the Evaluation of ZEB0 Questionnaire (unless otherwise specified). To study how schools used ZEB0, which factors influenced the use of ZEB0, and what the effects of ZEB0 were, a questionnaire was developed. The items in the questionnaire were designed to study the groups of factors in the theoretical framework: characteristics of ZEB0 as perceived by its user (A), implementation process features (B), school organisational characteristics (C), ZEB0 use (D), and the effects of ZEB0 use (E).
- § ZEB0 use: the use of the ZEB0 output as studied by means of the Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006 and/or by means of the interviews conducted in 2004, 2005, and 2006.
- § Conceptual and instrumental use: to study how schools used the ZEB0 output, a distinction was made between conceptual use and instrumental use of evaluation findings. Instrumental use was defined as the direct use of the ZEB0 output: decisions and actions are based on the output. Conceptual use refers to the indirect use of the ZEB0 output which may influence thinking about issues in a more general way, and in the longer term, may have an impact on the users' actions.
- § Evaluation: refers to the evaluation of the use of ZEB0 (unless otherwise specified).
- § Cohort 1 and cohort 2: To analyse the effect of the use of ZEB0 on pupil achievement, two cohorts of pupils were followed. Pupils in cohort 1 were followed from grade 3 (age 6) to grade 7 (age 11). Pupils in cohort 2 were followed from grade 4 (age 7) to grade 8 (age 12). Pupils' spelling attainment and mathematics attainment were tested on up to seven occasions: twice a year from 2002 (grade 3/4) to 2006 (grade 7/8).
- § LVS Mathematic (maths) and spelling (SVS) tests: the average pupil achievement level of schools was measured by means of spelling and mathematics tests from the pupil monitoring system (LVS) developed by Cito (the Dutch Testing and Measurement Institute). Cito developed the LVS to monitor pupil achievement in primary schools (age 4-12) over time.

Chapter 1

Self-evaluation in Dutch Primary Schools: Introduction and Research Questions

1.1 Introduction

School self-evaluation, quality care, and school improvement are important themes in current educational policy-making and are receiving increased attention in research. However, from the review of research on school self-evaluation by Kyriakides and Campbell (2004) it becomes clear that there is a need for more research into the use of and the effects of school self-evaluation systems. Visscher and Coe (2003) state that “although schools around the world want to use these systems we cannot be confident that they offer benefits to the schools, as they have yet to be rigorously evaluated”. A thorough evaluation of the use and the effects of various self-evaluation instruments is urgently needed.

In the following section (1.2), the concepts of school quality and school quality care are discussed and defined. These issues are closely related to the concept of school self-evaluation, presented in section 1.3. Next, in section 1.4 self-evaluation in general, in Dutch primary schools is discussed. This is followed in Section 1.5 by a description of the subject of the current study, ZEBO, a Dutch self-evaluation system for primary schools. Section 1.6 comprises the main aims of this study: the evaluation of the use and the effects of ZEBO. Finally, this chapter ends with section 1.7 an overview of this study.

1.2 School Quality and School Quality Care

Scheerens, Glas and Thomas (2003) state that educational quality can be defined on the basis of the outputs of the school. Based on the typology of effectiveness models developed by Quinn and Rohrbaugh (1983, in Scheerens et al., 2003), Scheerens et al. describe four models defining the criteria of organisational effectiveness:

1. The rational goal model: the central criteria for judging the organisational output are productivity and efficiency. Output in the case of schools may be defined in terms of average attainment level of pupils adjusted for prior achievement and other pupil intake characteristics (value-added pupil achievement);
2. The human relations model: human resource development is the central criterion for judging the organisational output. Work satisfaction and motivation of teachers are the terms used to define school output;
3. The open system model: in this model, the organisational output criteria comprise of growth and resource acquisition. This model emphasises the responsiveness of schools to environmental demands. Schools may create effective buffers against

external threats and they may manipulate their environment in order to safeguard and improve their own functioning. Growth (enrolment figures) and resource acquisition, (e.g., measured by the state of the buildings and equipment), are the criteria for judging the extent of responsiveness;

4. The internal process model: The organisational effectiveness criteria in this model are stability and control. The criteria for judging stability and control comprise of attendance rates, the number of teaching periods not given, and figures about the continuity in staffing.

Which criteria may be used for judging the quality of education in the current study? Quality depends on the point of view of the actors. Scheerens (1999) states that in a pluralistic and relativistic view, adherence to one of the criteria may be dependent on the actor's position regarding the organisation or the actor's organisation-theoretical preference. Parents probably have, for example, a different definition of school quality to the government or the schools' inspectors, and different parents will maintain different definitions. Some parents will emphasise the emotional well-being of their children, where other parents stress high achievement as the most essential determinant of quality (Deckers & Jacobs, 1994; Van Petegem, 1998a).

Although quality is to a certain degree dependent on the actor's perspective, it is necessary for the current study to arrive at a quantifiable definition of quality. Scheerens et al. (2003) point out that educational quality may easily be aligned with economic constructs such as productivity and efficiency (as in the rational goal model), which may be judged by value-added pupil achievement levels. Furthermore, Scheerens (1999) states that the effectiveness criteria described may also be ordered as means-goal relationships, whereby productivity is seen as the ultimate effectiveness criterion, as is displayed in Figure 1.1.

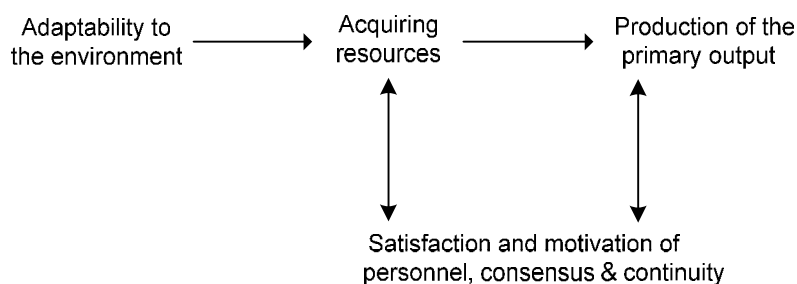


Figure 1.1 Means-goal relationships between effectiveness criteria (source: Scheerens, 1999)

For the current research Scheerens' definition is used, in which the school quality is defined as the average pupil achievement of a school adjusted for relevant pupil background characteristics (value-added pupil achievement). The achievement level of the school is influenced by criteria such as job satisfaction, availability of resources, and consensus among staff. In most school effectiveness research the output variable of

education is also defined in terms of value-added pupil achievement (Ledoux, Overmaat & Koopman, 1997; Hoeben, 1995; Van Petegem, 2001; Scheerens, 1997; Scheerens & Bosker, 1997; Scheerens, 1999; Scheerens, et al., 2003).

Quality care is a term describing the active focus of schools in ensuring the quality of their education, and, if possible, improving that quality (Visscher, 2002b; Hendriks, 2001). Quality care consists of *quality control* and *quality improvement*. To be able to improve the quality, the goals and the mission of the organisation (the target situation) must be clear, so that it is possible to determine the discrepancy between the current and the target situation. The process of gathering information on the discrepancy between the current and target situation is called *quality control*. *Quality improvement* refers to a situation in which a discrepancy is detected and this information must lead to actions which decrease that discrepancy (Doolaard & Karstanje, 2001). Quality care is a cyclical process (Deckers & Jacobs, 1994; Hendriks & Bosker, 2003), and originates in the Plan Do Check Act (PDCA)-cycle of Deming:

1. Plan: determining the goals of the organisation, the means that may be used to reach these goals, and the standards by which the realisation of these goals may be measured;
2. Do: execution of the quality policy. Agree upon responsibility, time scale, and instruments;
3. Check: evaluating whether or not the set quality goals are reached;
4. Act: decision-making for maintaining and improving the quality (in Hendriks & Bosker, 2003).

In the “Plan” phase, the initiative for quality care must be taken. One or more people must do the preparation. Next, consensus is needed in the team about the necessary actions that need to be carried out for quality care. In the “Do” phase, data must be collected on the functioning and quality of the school. The data must be analysed, and after analysis, priorities must be fixed. Which areas require improvement? How can this be realized? In the “Act” phase, improvement plans must be implemented. The “Check” phase is the actual evaluation phase. Have the goals been reached? Have the improvement plans been successfully implemented? If this is the case, it must be ensured that the school does not revert to old methods. If it is doubted that the implementation was successful, the schools must go back to the “Do” phase and the quality care cycle starts again (Hendriks & Bosker, 2003). Self-evaluation forms an important aspect of this quality care cycle, as will be explained in the next section.

1.3 School Self-Evaluation

School self-evaluation is closely related to such concepts as quality, quality care and quality control. School evaluation is described by Scheerens et al. (2003) as judging the value of schools on the basis of systematic information gathering, in order to support decision making and learning. School self-evaluation, in turn, is defined as school evaluation where school staff carries out the evaluation regarding their own school.

Van Petegem (2001) gives a similar definition, but adds that the information gathered should be used for school improvement. He states that school self-evaluation can be described as a procedure started by the school for gaining information on the functioning of the school, and on the design and goals of education, for taking policy decisions on school improvement (Van Petegem, 2001).

Combining these two descriptions and for the purposes of the current research, school self-evaluation is defined as *a procedure involving systematic information gathering which is initiated by the school itself and aims to assess the functioning of the school and the attainment of its educational goals for the purposes of supporting decision-making and learning and for fostering school improvement as a whole*. Self-evaluation is part of the quality care cycle that was discussed above, and it is mainly aimed at determining the quality, and improving the quality of education (the Check phase).

School self-evaluation systems have been introduced into schools around the world for several reasons. First of all, decentralisation has taken place in many countries. In the Netherlands, for example, schools are responsible for the quality of their education. This means that schools must evaluate their functioning on a regular basis to assess, maintain, and, if necessary, improve their quality (Hendriks, Doolaard & Bosker, 2002).

Moreover, a political climate of public sector accountability has arisen. Schools are faced with public judgments of their effectiveness. However, this public performance information frequently includes average, raw pupil achievement scores, and does not include value-added scores. Schools are therefore in need of more accurate and reliable information about their performance in order to make sound decisions regarding whether or not improvement is necessary (Coe & Visscher, 2002a).

Schools are more independent now than they were in the past and have the opportunity to become more attractive to potential pupils. They must distinguish themselves from their competitors more explicitly (Marx, De Vries, Veenman & Slegers, 1995). The improvement of education may be necessary in order to compete with other schools (Deckers & Jacobs, 1994; Marx et al., 1995).

Instruments for self-evaluation can help schools in these matters. Self-evaluation may offer a starting point for further analysis and assist in the diagnosis of specific points in the school's functioning. It may also be a useful way to inform relevant audiences about the school's quality (Hendriks et al., 2002).

Several instruments for self-evaluation are available. These instruments help schools to collect data on educational indicators. Data on these indicators help schools to determine their quality. Scheerens (1991) defines educational indicators as statistics that allow for value judgments to be made about the most important aspects of the functioning of educational systems. Usually five kinds of educational indicators are distinguished (Hoeben, 1995; Porter, 1991; Scheerens, 1991; Scheerens et al., 2003):

1. Input indicators: for example, the characteristics of the pupil population, the size of the school, the composition of the pupil population, the constitution of the teachers' team, the financial and human resources available to the school. It is not very likely that schools and teachers are able to influence these indicators;
2. Process indicators: such as organisational characteristics, the school plan, the goals, the education offered, the learning environment, educational leadership, the time spent on tasks, homework, evaluation frequency, absenteeism, vandalism, and absence due to illness. These factors involve the provision of education by schools and teachers;
3. Output indicators: factors such as success rates of pupils, exam results, achievement, attitudes, and value-added pupil achievement results;
4. Impact indicators: these factors refer to changes in other sectors of society which may be seen as the effects of education, such as the impact of education on youth unemployment and delinquency rates;
5. Context indicators: factors which refer to society at large and structural characteristics of national education systems, such as demographics and the structure of schools in the country.

Information on process indicators is most suited for determining the quality of education. Process indicators may offer possible explanations about why certain schools perform better than others. These indicators, in general refer to characteristics of schools which may be manipulated (Scheerens, 1991). In the literature on school effectiveness (Hoeben, 1995; Bosker, 2001; Hendriks, 2001; Porter, 1991; Griffith, 2002; Scheerens, 1991) several process indicators are frequently mentioned as being important for school quality, such as strong educational leadership, monitoring of outcomes, and regular evaluation of pupil progress.

Information on these indicators, obtained by using a self-evaluation instrument, makes it possible to detect problems quickly and to devise potential solutions, if a quality problem exists (Hoeben, 1995). Schools may use the information on these indicators to improve the quality of education, which is the most important function of school self-evaluation (Saunders, 1999, Olthof, Emmerik & Troost, 1993; Visscher, 2002). Several schools in the Netherlands are using these kinds of self-evaluation instruments.

1.4 Self-Evaluation in Dutch Primary Schools

Dutch schools traditionally have considerable autonomy. They have always been free to choose the religious, ideological and pedagogical principles on which they base their education, as well as how they choose to organise their teaching activities (Ministerie van Onderwijs, Cultuur & Wetenschappen, 1999). This freedom has led to a situation where both public and private schools are funded equally by the Government. Since the 1980s the process of further decentralising competencies from the national level to the level of schools and municipalities has been initiated, schools have received more autonomy regarding their administration and finances; some other tasks have been decentralised to the municipalities (Hendriks et al., 2002).

Since August 1998, the Dutch “Quality Law” stipulates that schools are responsible for the quality of the education they provide and for pursuing policies that ensure improvement. The law also prescribes that all schools must develop a quality assurance system.

In keeping with the aforementioned developments and based on two memoranda from the Dutch Ministry of Education, ‘Diversity and Guarantee’ (1999) and ‘Towards Stimulating Supervision’ (2000), the Dutch Schools Inspectorate has developed two new types of school supervision: Integral School Supervision (IST) and from 1999 Regular School Supervision (RST) (Van Bruggen & Mertens, 2001; van der Grift, 2001). As from September 1st, 2002, when the new law on the Supervision of Education went into effect, the new role of the Inspectorate was also laid down in law. 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, provided they fulfil the standards set by the Inspectorate. The Inspectorate is entitled to supervise schools on more quality aspects than before. The Inspectorate examines the schools’ results and a number of key quality characteristics relevant to the teaching-learning process in regular supervision. The Inspectorate has a quality control task and also evaluates the school plan and the prospectus. The school plan contains the school’s policy on the quality of education and school improvement activities planned for the next four years. The school plan is an integral internal policy document as well as an accountability document for the Inspectorate. The school prospectus gives information on a school’s objectives, its educational activities and the results achieved. The school prospectus is a public record for parents and pupils (Hendriks et al., 2002). From September 1st, 2002, the Inspectorate is also authorised to promote the quality of the school (Inspectie van het Onderwijs, 2002; Ministerie van Onderwijs, Cultuur & Wetenschappen, 2000-2002; Renkema, 2002).

During regular supervision the Inspectorate examines whether there are any problems in the school. If a problem exists, integral supervision is carried out. With integral supervision the Inspectorate examines whether improvement is needed. If improvement is deemed

necessary, the school is obliged to develop a plan of action and to implement it (Hendriks et al., 2002; Hendriks, 2001, Inspectie van het Onderwijs, 2002; Hoeben, 1995).

School self-evaluation is not compulsory in the Netherlands although highly recommended by the Inspectorate. Schools are legally required to have a school policy on maintaining and improving the quality of education. Furthermore, the Supervision of Education law from 2002 states that “external inspection can be more restricted in a school with a more sophisticated quality care system”. For these reasons, an increasing number of schools are starting to implement a form of school self-evaluation. More than 70 different instruments for school self-evaluation are available (The Standing International Conference of Central and General Inspectorates of Education, 2003).

1.5 ZEBO

A study of the school self-evaluation instruments used in the Netherlands revealed that the aims of the various instruments differed. Some instruments describe the current functioning of the school, others are aimed at improving the functioning of the school, maintaining its current functioning, or further developing the school (Cremers-Van Wees, Rekveld, Brandsma & Bosker, 1995).

However, a study of school self-evaluation instruments indicates the presence of technical weaknesses in the available instruments, such as a lack of attention to the reliability and validity of the instruments (Cremers-Van Wees et al., 1995).

ZEBO (in Dutch the acronym stands for Self-Evaluation in Primary Schools: ZelfEvaluatie in het BasisOnderwijs) has been developed as a response to this situation. ZEBO is a self-evaluation instrument for Dutch primary schools which took five years to develop. The process started with a literature review of research on school effectiveness, school improvement and performance indicators. Thirteen process variables, frequently mentioned in school effectiveness research were selected for the development of ZEBO (Scheerens & Bosker, 1997). The factors and components, found on the basis of the school effectiveness research, were presented to principals and teachers. Their opinions and the conclusions of the Committee for Primary Education Evaluation, in 1994, were the basis of an initial selection of process variables.

After the selection of the process variables, an inventory of all available instruments and questionnaires for self-evaluation was compiled. Next, instruments were selected or constructed to measure these variables. Process variables were selected at school level and at classroom level. Table 1.1 displays an overview of the selected process variables and the level of data collection. The selected variables are all associated with high achievement in the school effectiveness literature (Hendriks et al., 2002).

Table 1.1 *Process variables selected and level of data collection (source: Hendriks et al., 2002, p. 124)*

Process variables	Information collected at the level of		
	<i>school management</i>	<i>teachers</i>	<i>pupils</i>
School level			
- Achievement orientation/ high expectations	X	X	X
- Educational leadership	X	X	
- Staff professional development	X	X	
- Pupil care; measures that enable adaptive education	X	X	
Consensus and cohesion among staff:			
- Frequency and content of formal staff meetings with school management	X	X	
- Frequency and content of informal meetings among teachers (co-operation)	X	X	
School climate:			
- Relationships amongst staff	X	X	
- The relationship between school management and staff	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 amongst pupils			X
- Teacher support and relationships between teacher and pupil			X

In 1997 and 1998, two trials of ZEBO took place in 43 schools and 58 schools respectively, in the Dutch Twente region. In 1999, the final field test took place in a representative sample of 123 schools in the Netherlands. ZEBO was slightly modified after each development phase on the basis of the analyses of reliability and validity. However, the instrumentation basically remained the same. The final market version of ZEBO was released in a computerized form in 2003. Now schools may use ZEBO whenever they need the information, and the output is immediately obtainable (Hendriks et al., 2002).

Questionnaires are used to measure the process variables for (Hendriks & Bosker, 2003):

§ School management and teachers: the management provides information at school level. Management Topics: co-operation and consultation, pupil care, working environment, educational leadership, professional development of staff, and agreement on goals and expectations. Teachers provide information at school level and at classroom level. The first nine scales deal with variables at the school level, almost identical to the nine scales of the school management questionnaire. The last four scales contain variables at the classroom level. Teacher Topics: structured education, adaptive education, classroom climate and learning time;

- § Pupils from Grades 4-8 (aged seven-twelve years): provide information at classroom level. Topics: structured education, adaptive education, classroom climate, and learning time;
- § Pupils from Grade 3 (aged six-seven years): also provide information at classroom level. This questionnaire is a modified version of the questionnaire for pupils from grades 4-8. The questionnaire differs in formulation, possible answers and in the way in which the questionnaire is administered. The pupils are guided through the questionnaire by a teacher (other than their own), a coach, or another adult. Pictograms are used in the questionnaire and pupils may only answer 'yes' or 'no'.

Schools may choose which (sections) of the questionnaires they wish to use. They may, for example, use only the Pupil Questionnaire, if they need information regarding a specific issue. They might use only the questionnaires for the upper grades, or they may choose not to use Pupil Questionnaires at all, and only use the questionnaires for the principal and the teachers.

After filling out the questionnaires in the schools, two kinds of output are available. (Hendriks et al., 2002):

- § A school report which contains a graphic and a written representation of the results for each scale for the school in comparison with the schools from a total national sample. The results of the teachers are also compared with the results of the school management;
- § A classroom report which contains a graphic representation and a textual explanation of the results of the Pupil and Teacher Questionnaires. The results of pupils of a certain grade are compared to the results of pupils in the national sample from that same grade. The results of the pupils are also compared with the results of the teachers.

Schools have options about what kind of output they want to generate. Some schools might want to generate only the global output for some scales, while others might want to analyse item scores or even individual scores. The reports identify the significant differences between the results of the teacher and the school management, or between pupils and the classroom teacher. However, the reports do not identify the direction of the differences. These differences may form a starting point for internal debate and consultation (Hendriks et al., 2002). 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 & Bosker, 2003; Hendriks, 2001).

1.6 Research Questions

The first goal of this study is to systematically acquire detailed knowledge on the use of the school self-evaluation instrument, ZEBO, within schools. According to Coe and Visscher (2002a) little is known about the nature and extent of the use of self-evaluation instruments within schools. In summary, the first research objective investigates how and to what extent schools use the output generated by ZEBO.

Since the use of a self-evaluation instrument should lead to school improvement, the second goal of this study involves researching the effects of the use of ZEBO. It is known that improvement of school performance, in terms of average pupil achievement levels, takes many years. Hence, to what degree improvement on important prerequisites for improved school performance, (such as improvement on process indicators, for example, teaching, communication, and educational leadership) may be observed, will also be investigated. Finally, the use of a self-evaluation instrument may also have unintended consequences, for example, increased workload or stress. The prevalence of these negative, unintended effects will also be studied. In summary, the second research objective investigates the (intended and unintended) effects of the use of ZEBO.

In several studies (Weiss, 1998a; 1998b; Coe & Visscher, 2002b; Van Petegem & Vanhoof, 2002a; 2002b; 2004) it was observed that evaluation findings were under-utilised. Valuable new information was often found not to lead to improvement-oriented behaviour. Social support, availability of additional resources, and a strong motivation to improve were found to be important preconditions. Moreover, evaluation use may have been obstructed in various ways, for example, the evaluation results were not disseminated within the school; school staff did not understand or believe the results; or school staff did not know how or did not have the required skills to improve the results (e.g. how to improve pupil achievement). Effective quality assurance requires knowledge and skills, such as statistical knowledge, which may not be available among school staff. However, there is still a dearth of systematic knowledge available on which factors may help schools to make effective use of self-evaluation instruments and data. Therefore another important goal of this study is to systematically acquire insight into the factors that are decisive for the successful use of self-evaluation instruments. In summary, the third research objective investigates which factors influence the use of ZEBO.

In summary, the main research questions enquire:

1. How and to what extent do schools use ZEBO?
2. What are the effects of the use of ZEBO?
3. Which factors influence the use of ZEBO?

1.7 Overview of this Study

Chapter 2 provides a detailed description of the theoretical background. The research method is presented in chapter 3, along with descriptions of the various research instruments used. The empirical results appear in chapters 4 to 6. The focus of chapter 4 is on the results regarding the use of ZEBO. Chapter 5 presents the effects of the use of ZEBO, and chapter 6 explains which factors influence the use of ZEBO. In chapter 7, conclusions are drawn on the overall results of this study, along with the theoretical and practical implications of those results.

Chapter 2

Theoretical Framework

2.1 School Performance Feedback Systems

A theoretical framework for School Performance Feedback Systems (SPFS) developed by Visscher (2002) was applied in this study to systematically acquire detailed knowledge on the use and effects of ZEBO. According to Coe and Visscher (2002a), 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).

ZEBO is one of those systems (it is however a school internal system) which can provide schools with confidential information (in the form of school and classroom reports) on their performance and functioning as a basis for school self-evaluation. In the next section, the theoretical framework for studying ZEBO is presented.

2.1.1 A Theoretical Framework for Studying School Performance Feedback Systems

Visscher (2002) developed a framework to study SPFSs. Figure 2.1 displays this framework with assumed relationships between five groups of factors: the characteristics of the SPFS (block A), the implementation process features (block B), the school organisational characteristics (block C), the use of the SPFS (block D), and the effects of its use (block E). Each block includes a multitude of variables.

The characteristics of the SPFS are assumed to influence the use of the SPFS, which leads to intended and unintended effects. The implementation process features influence both SPFS use and the school organisational characteristics, which again influence SPFS use (Visscher, 2002).

The theoretical framework presented in Figure 2.1 will be used as a starting point for this study. As stated in the previous chapter, three main questions form the basis of this study:

1. How and to what extent do schools use ZEBO?
2. What are the effects of the use of ZEBO?
3. Which factors influence the use of ZEBO?

Educational system

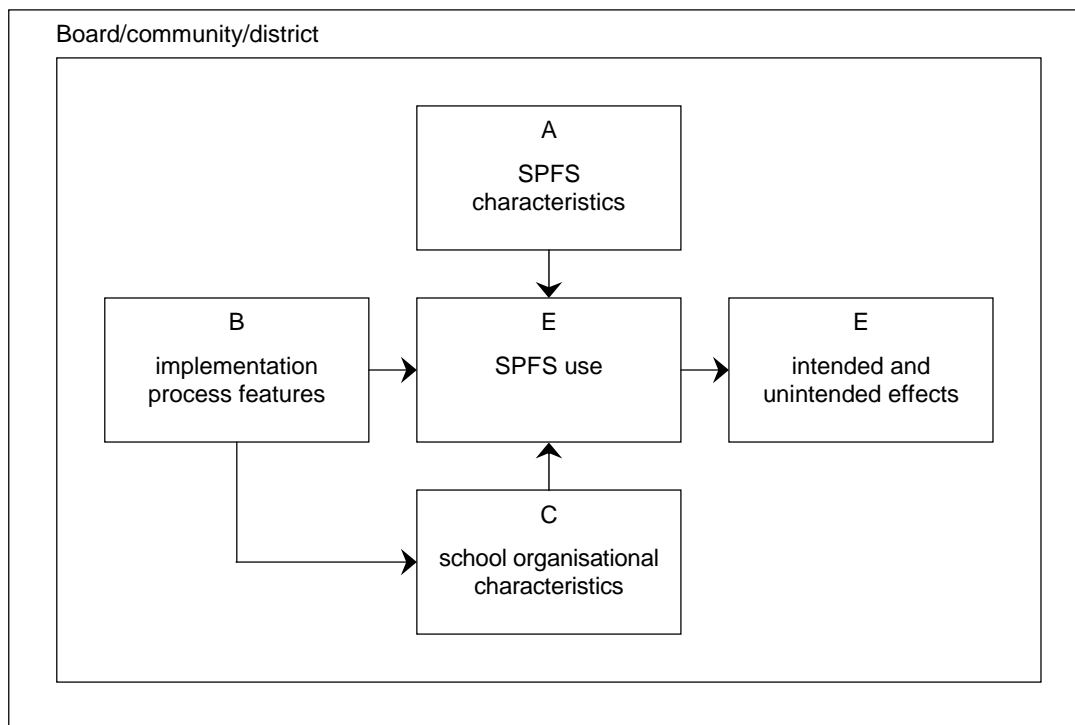


Figure 2.1 The relationships between the groups of factors (Based on Visscher, 2002, p. 43)

2.2 The Use of School Performance Feedback Systems

The first focus of this study is to investigate the use ZEBOS. The use of ZEBOS 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 ZEBOS 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. Leithwood, Aitken and Jantzi (2001) state that although enlightenment does not necessarily directly lead to action to improve the quality of education, it is expected that such changes in thinking may eventually influence the actions of school staff.

The ZEBOS 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 ZEBOS 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, 2002b).

The following variables regarding the use of the ZEBO output are included in this study:

- § D1 Study of output;
- § D2 Resulting measures;
- § D3 Discussion of output;
- § D4 Resulting new insights;
- § D5 Highlighted problems;
- § D6a Solutions by school staff;
- § D6b Solutions by individuals;
- § D7a Measures by school staff;
- § D7b Measures by individuals.

Following Weiss (1998), a distinction is made between instrumental use and conceptual use of the ZEBO output. Rossi, Freeman and Lipsey (1999) define instrumental use as the direct use of evaluation findings: the results are analysed 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 utilised to study how schools use ZEBO.

2.3 Effects of School Performance Feedback Systems

The goal of using a SPFS should be the monitoring and improvement of school performance. As improving school performance in terms of improved average pupil achievement levels takes a long time, it is not likely that ZEBO will have a measurable impact on pupil achievement during the first years of its use. However, some studies into the use of SPFSs do indicate small positive effects on the quality of education. Tymms and Albone (2002), for example, report a very small positive effect of Performance Indicators in Primary Schools (PIPS) on pupil achievement. Thus, it was deemed worthwhile to study the effect of ZEBO on pupil achievement.

Next, the degree to which observation can be made of specific prerequisites for improved school performance, in terms of pupil achievement, was investigated. Various studies into the effects of the use of a SPFS (Teddlie, Kochan & Taylor, 2002; Hendriks et al., 2002; Gray, 2002; Rowe, Turner & Lane, 2002; Tymms and Albone, 2002) indicate some positive effects on specific prerequisites for improved school performance, such as an effect on teaching behaviour (Davies & Rudd, 2002; Webb, Vulliamy, Häkkinen, & Hämäläinen, 1998).

To determine other possible effects of ZEBO, 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 ZEBO is to improve the quality of the school, eventually in terms of pupil achievement. Therefore, the effects of ZEBO on those variables which seem to be positively associated with pupil achievement were included in this research.

Firstly, the effect of ZEBO on the amount of consultation on the schools' functioning and quality was studied. Without consultation and communication it is very hard, if not impossible, to induce any change in an organisation (Scheerens and Bosker, 1997).

Secondly, several school effectiveness studies have highlighted educational leadership as a characteristic that is positively associated with pupil achievement (Scheerens, 1990; 1991; Mortimore, 1998; Reynolds, Hopkins, Potter & Chapman, 2002; Scheerens & Bosker, 1997; Witziers, Bosker, & Krüger, 2003; Leithwood, Jantzi & Steinbach, 1999). ZEBO provides a school with information which may be used to improve the functioning of the principal. Hence, the effect of ZEBO on educational leadership was studied.

Furthermore, the use of a SPFS may affect the 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. Reynolds et al. (2002) found in their literature review that commitment to staff development and training is positively associated with pupil achievement. The professional development of teachers may be promoted by making courses available to teachers, and by offering guidance. (Verloop, 1995; Van den Berg & Vandenberghe, 1999). ZEBO provides schools with information which may be used to enhance the professional development of their teachers. Based on the ZEBO output school staff may conclude, for example, that more in-service training is required or that (newly qualified) teachers need coaching.

Another aspect of ZEBO which was studied is its possible effect on the achievement orientation of school staff. A school staff with a high pupil achievement orientation is an important prerequisite for improved school performance (Visscher, 2002; Scheerens & Bosker, 1997; Reynolds et al., 2002). According to Scheerens and Bosker (1997), a strong achievement orientation includes a clear focus on the mastery of basic subjects, fostering high expectations of pupil achievement, and the use of records to monitor pupil progress. ZEBO provides schools with information about their achievement orientation: the degree to which learning profits are central within the school, and a school aims at obtaining high pupil achievement results, the extent to which high expectations of pupils are held (at school level), and the degree to which the school's goals are clear. Schools may use this information to move towards a stronger achievement-oriented policy.

Team cohesion may be influenced by the use of ZEBO. The degree of team cohesion is a school characteristic also consistently associated with high pupil achievement (Scheerens, 1990; Scheerens & Bosker, 1997; Reynolds et al., 2002).

Moreover, ZEBO use may influence 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. It may encompass a wide variety of issues including health, social and moral education, behaviour management and emotional support. 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.

Improved teaching 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). ZEBO provides schools with information on the behaviour of teachers as perceived by their pupils. Teachers may use this information to improve their didactic methods.

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). Verloop and Van der Schoot (1999) state that learning processes in general will improve if the person who is learning receives frequent information on his or her progress and if the teacher is able to diagnose possible problems and failure to keep up with other pupils.

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 & Vandenberghe, 1999). Because pupils learn in different ways and at varying rates, the instruction, the subject content, and the time pupils have to complete a task should take account of these differences. ZEBO provides schools with information, at class level, on the degree to which teachers take differences between pupils into account.

Finally, the use of ZEBO may lead to certain negative effects. The (administrative) workload of teachers and principals may increase as a result of using a self-evaluation instrument (Visscher, 2002; Van Petegem, 1998a). Moreover, participants may feel threatened by the evaluation, and evaluations may evoke defensiveness (Clift, Nuttall, & McCormick, 1987). Another unintended consequence may be the pursuit of short term targets at the expense of legitimate long term objectives (Smith, 1995). Finally, school evaluation may have a de-motivating impact on teachers, especially in poorly performing schools (Van Petegem, Vanhoof, Daems, & Mahieu 2005).

Based on the above, the following variables were selected for in-depth study:

- § E2.1 Consultation on school functioning and quality;
- § E2.2 Educational leadership;
- § E2.3 Professional development;
- § E2.4 Achievement orientation;
- § E2.5 Team cohesion;
- § E2.6 Pupil care;
- § E2.7 Didactic methods;
- § E2.8 Pupil achievement evaluation;
- § E2.9 Adaptive education;
- § E3.0 Negative effects.

2.4 Factors Influencing the Use of School Performance Feedback Systems

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

Characteristics of ZEBO (A)

SPFSs may differ in the degree to which the output they provide is perceived as relevant by their users (Visscher, 2002). Rowe et al. (2002) studied the use of performance feedback in the Year 12 Victorian Certificate of Education (VCE) assessment program. They state that the provision of accurate, appropriately-adjusted and responsibly-presented performance data is very important. The data must be accessible to the users and presented in an appealing way. Teddlie et al. (2002) conclude similarly, 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 ZEBO. Fullan (1991) states that the people involved must feel the need for the innovation. School staff must trust the output. Teddlie et al. (2002) state in this context that if systems are perceived as attacking school staff, schools are unlikely to respond positively and accept the information.

The user-friendliness of the system (Visscher, 2002) is also assumed to be an important factor in the use of ZEBO. It should not be too difficult to use ZEBO 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, 2002b).

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

In short, the following variables concerning the characteristics of ZEBO (as judged by its users) were investigated:

- § A1 Relevance of output;
- § A2 timeliness of output;
- § A3 Accuracy of output;
- § A4 Fit of output with user needs;
- § A5 Ease of data entry;
- § A6 Ease of output generation;
- § A7 Ease of data alteration;
- § A8 Clarity of output;
- § A9 Time requirement of use;
- § A10 Ease of use.

Implementation Process Features (B)

Several authors stress the importance of implementation (Fullan, 1991; Visscher, 2002; Gray, 2002; Rowe, et al., 2002). As Fullan (1991) states “the proof is in the putting: how change is put into practice determines to a large extent how well it fares” (p. 9).

User training and support are assumed to play an important role in the successful implementation of a SPFS (Visscher, 2002). Users of ZEBO need certain skills and knowledge to successfully implement it, such as knowledge of how to interpret the statistics generated by ZEBO. 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 school receive (in their opinion) enough training and support, they are more likely to use ZEBO 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 organisational 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. Pressure may lead to action, which may lead to

improvement. 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.

Clift et al. (1987) studied several school self-evaluations in primary and secondary schools in England and Wales. They concluded that for the self-evaluation to be successful, it is very important that its purpose is clear. Often the goals are presented in phrases like “institutional development”, “professional accountability, and “school improvement”. These phrases look very good, but their meaning is often unclear to most teachers and also to most principals. 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.

The variables with regard to the implementation process features included in this study comprise:

- § B1a Hours of training and support received;
- § B1b Satisfaction with amount of training;
- § B2 Satisfaction with content of training;
- § B3 Satisfaction with amount of support;
- § B4 Satisfaction with content of support;
- § B5 Encouragement by principal;
- § B6 Pressure to implement;
- § B7 Clarity of goal.

School Organisational 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 (Teddlie et al. 2002; Rowe 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). Van Den Berg, Vandenberghe and Slegers (1999) stress the importance of the feelings of the teachers with regard to innovations. If teachers perceive an innovation as negative or unnecessary, this may result in an unsuccessful implementation of the innovation.

Furthermore, school self-evaluation requires that schools devote a substantial amount of time, energy and resources to it (Teddlie et al., 2002; Kimball, 2002; Davies & Rudd, 2001; Visscher, 2002). Schools are often very busy with routine activities that take up the most of their time. Consequently, they do not have much time to invest in innovations. 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 ZEBO.

Another school organisational 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 of schools as the competence of schools to implement innovations initiated by both the government and the school and to make sure that, if necessary, both kinds of innovations are related to each other. In short, the innovation capacity is the capacity of schools to implement innovations in a successful manner (Geijsel, 2001). Geijsel et al. (1999) identify four components of the innovation capacity of a school: participation in decision making, collaboration among teachers, transformational school leadership, and the functioning of the school as a learning organisation. Higher innovative schools are more likely to use ZEBO as they have transformational leaders, who will probably encourage the use of ZEBO. Transformational leaders are leaders who focus on the commitments and capacities of the organisation's members. According to Leithwood et al. (2000), higher levels of personal commitment to organisational goals and greater capacities for accomplishing those goals are assumed to result in extra effort and greater productivity. In addition, highly innovative schools are learning schools. According to Macbeath and Mortimore (2001), there is a commitment by staff to reflect, to adapt and to learn in these schools. School staff are not afraid to try something new and they are encouraged to experiment. Staff in highly innovative schools are involved in school decision making. So, it is likely that if teachers are involved in general school decision-making, they were also involved in the decision to participate in the ZEBO-project. Also, collaboration among teachers is important. Van Gennip and Claessen (1993), for example, conclude from their study into the differences between innovating schools and non-innovating schools that collaboration of staff plays an important role. Team cohesion and collaboration are important for successful innovations in schools.

Another school organisational characteristic which is considered relevant for use of a SPFS is actual score on the SPFS (Visscher, 2002). Relatively low scores combined with a pressure strategy may motivate schools to try to improve performance by using the results of a SPFS. A new ZEBO score was devised for the purposes of this study by combining several ZEBO output scores, it was decided to call this variable ZEBO score.

The following organisational variables, based on the above, were included in this study:

- § C1 The innovation attitude of staff;
- § C2 Time and resources for innovation activities;
- § C3 School innovation capacity;
- § C4 ZEBO score.

To answer the research questions in this study, a new framework, based on Visscher's framework, was developed. Figure 2.2 represents all the variables included in the framework. The variables in the framework have been selected based on the framework for studying School Performance Feedback Systems as developed by Visscher (2002), the contents of ZEBO, and findings from school effectiveness and school improvement research.

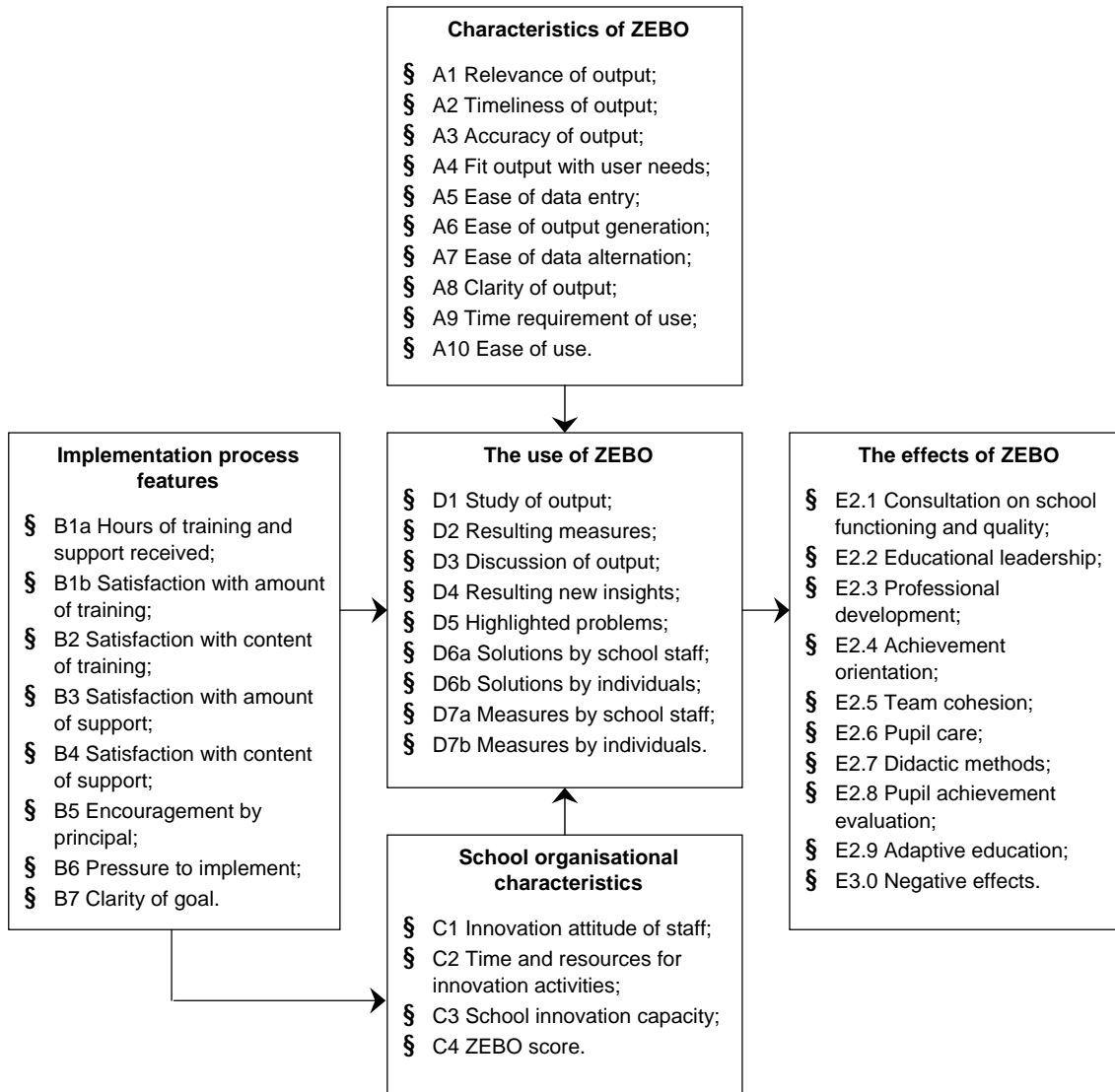


Figure 2.2 The factors expected to influence the use of ZEBO, and the studied effects of ZEBO use

Chapter 3

Method

3.1 Research Design

The first objective of this study was to evaluate how and to what extent schools use the school self-evaluation instrument ZEBO. Secondly, the effects of the use of ZEBO were investigated (the effects on pupil achievement, and other possible effects, such as improved educational leadership). The third objective was to study which factors influence the use of ZEBO. To answer these research questions, data collection started in the 2001-2002 school year and continued during the 2002-2003, 2003-2004, 2004-2005 and 2005-2006 school years. During these five years, the study followed two cohorts of pupils from grade 3 (age 6) and grade 4 (age 7) to respectively grade 7 (age 11) and grade 8 (age 12). This chapter will start by presenting information on the sample (3.1.1), and on data collection and the instruments used (3.2). Then, information will be given on how the data were analysed (3.3).

3.1.1 Sample

The target population in this study consisted of all the primary schools in the Netherlands, with the exception of special primary schools. A purposive sample of primary schools was drawn. All 312 schools in the district of the school advisory service, Expertis, were asked to participate in the study.

Seventy-nine Dutch primary schools were willing to participate. Whether the sample was representative for Dutch primary schools was tested. The sample was found not to be representative with respect to school denomination. Analysis showed that the sample included more public schools and fewer protestant schools than the population. The number of Catholic schools was representative for the population (Table 3.1).

Table 3.1 *Denomination of schools*

Denomination	Population	Sample
Public	33.5%	50.7%
Catholic	29.5%	30.7%
Protestant	29.8%	16%
Other	7.2%	2.6%
Total	100%	100%

By investigating school size and pupil population, it was estimated whether the schools were representative of the population using a Levene's test for equality of variances and a test for equality of means. Both analyses showed that the sample was not representative for school size ($F=10.61$, $p=0.01$ and $T=3.98$, $P=0.00$). The schools in the sample had a smaller average school size. The sample was representative regarding the composition of the pupil population of schools ($F=0.26$, $p=0.61$ and $T= -0.42$, $p=0.67$) in terms of the social economic status of the parents of these pupils.

It must be taken into account that ZEBO was studied in only one region of the Netherlands, and consequently, the number of respondents participating in the study was limited. The sample was found not to be representative for school size and denomination. Therefore it was not possible to generalise the results to the entire population.

A cohort of regular pupils from the schools which participated in the study was followed from grade 3 and 4 (school year 2001/2002, age 6 and 7) to grade 7 and 8 (school year 2005/2006, age 11 and 12). Regular pupils are pupils from the same cohort who start education in group 3 or 4 and follow the same group course. This excludes those pupils who leave school and the non-promoted pupils. According to Moelands et al. (2000), in order to be able to generate conclusions on the quality of education, based on results of groups of pupils on regular tests, it is important to follow pupil groups with a constant composition. Several pupils left the cohort during the five years of this study for several reasons. A record of the reasons for pupil attrition was kept, to ensure that schools did not leave out the lowest scoring pupils.

Within the 79 schools which participated, 158 classes and 3.220 pupils (1.591 pupils from grade 3 and 1.629 pupils from grade 4) participated in the school year 2001-2002. During this year seven schools stopped participating. One school experienced severe technical problems with the ZEBO system, another was too busy relocating and five schools chose to continue with other quality care instruments (either voluntary or obliged by the school board). This brought the total number of schools participating in the research to 72 (2.542 pupils) during the second evaluation. Before the third evaluation another thirteen schools stopped using ZEBO. These schools chose to continue with another quality care instrument. Of those schools, four schools resided under one school board which had made the decision to continue with another instrument. The study started with 3.220 pupils and ended with 2.431 pupils; the causes for the decrease in participation included schools which stopped participating in ZEBO use, pupils changing schools, repeating grades and referral of pupils to special education.

In Table 3.2, the characteristics of the cohort of pupils are given (for school year 2001/2002 and for school year 2005/2006). As may be seen, almost as many girls as boys participated in the study, and most pupils have highly educated parents. Only a few Dutch and ethnic minority pupils with poorly educated parents participated in this study.

Table 3.2 Characteristics of pupils participating in the first test administration (2001/2002) and in the last test administration (2005/2006)

	Number of pupils at the start of this study 2002 (%)	Number of pupils at the end of this study 2006 (%)
Boy	1290 (50.1)	1233 (50.7)
Girl	1252 (48.6)	1198 (49.3)
<i>Socio-economic-status (SES)</i>		
Children with highly educated parents	1884 (73.1)	1823 (75.0)
Dutch children of parents who have had little education	404 (15.7)	387 (15.9)
Ethnic minority children of parents who have had little education	221 (8.6)	221 (9.1)

3.2 Data Collection and Instruments

Various instruments were used to collect data in this study:

- § The school self-evaluation instrument ZEBO;
- § Standardised spelling and mathematics pupil achievement tests to assess pupil achievement (LVS);
- § A pupil form to gather background information about the pupils;
- § The teacher and principal questionnaire on the use of ZEBO (Evaluation of ZEBO Questionnaire);
- § Interviews with teachers and principals about the use of ZEBO.

Appendix 3.1 shows the data collection timeframe. Quantitative and qualitative research methods were both employed in this study. Quantitative methods (Evaluation of ZEBO Questionnaire, ZEBO output and pupil scores on pupil achievement tests) were employed to produce data which is possible to generalise to some larger population. Qualitative data (interviews) were collected to generate detailed process data (Steckler, McLeroy, Goodman, Bird, & McCormick, 1992). More detailed information on the instruments used for data collection will be presented next.

3.2.1 ZEBO

Schools used the computerized version of ZEBO for the first time in 2002 or (due to technical problems) in 2003. Schools used ZEBO for the second time in 2004, and for the third time in 2006. Schools were free to use ZEBO more often if they wished. Table 3.3 presents the number of schools which used ZEBO in each phase of the study. During each phase, several schools did not use ZEBO for the following reasons:

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

Table 3.3 *Number of schools using ZEBO and unable to use ZEBO and reason given, per school year*

	2002/2003	2004	2006
No. of schools using ZEBO	64	58	43
Reason given for not using ZEBO:			
Technical problems	4	3	
School renovation			1
Management problems or changes	4	4	3
Change of school self-evaluation instrument	7	4	9
Chose to administer ZEBO later		3	12

The results of these ZEBO measurements were collected in order to study the score on the ZEBO scales. Whether or not the use of ZEBO is dependent on the results of the self-evaluation was investigated. If the ZEBO output is positive, there is probably no need, or a limited need to use the output. To investigate this, a new ZEBO score was devised for the purposes of this study, by combining several ZEBO output scores. It was decided to call this variable ZEBO score.

Reliability and Validity of the ZEBO Instrument

Almost all ZEBO scales met the criterion of reliability, expressed in internal consistency (Cronbach's alpha) both at the individual and aggregated levels (Table 3.4). Conclusions on the validity of the ZEBO instrument were mixed. Several variables correlate with each other, so it may be expected that these variables measure similar concepts to a certain extent (Hendriks, Doolaard, & Bosker, 2002). For further information on validity and reliability, the reader is referred to Bosker and Hendriks (1997).

Table 3.4 *Reliability of scales in measuring process variables (source: Hendriks, Doolaard & Bosker, 2002, p. 126)*

School level	Reliability at individual (school/head/teacher) level		Reliability at aggregated school level
- Achievement orientation/high expectations	0.8		0.7
- Educational leadership	0.8		0.5
- Staff development	0.8		0.8
- Pupil care; measures which enable inclusive education	0.8		0.6
Consensus and cohesion among staff:			
- Frequency and content of formal staff meetings with school management	0.8		0.7
- Frequency and content of informal meetings among teachers (cooperation)	0.8		0.6
School climate:			
- Relationships between staff	0.9		0.9
- Relationship: the role of school management	0.9		0.7
- Workload	0.8		0.6
Classroom level (grade 4 – 8)	Individual pupil level	Individual teacher level	Aggregated (classroom) level
- Achievement orientation/high expectations	0.7	0.8	0.8
- Structured instruction	0.6		0.8
- Adaptive instruction	0.7		0.5
- Time on task	0.8		0.4
- Classroom climate	0.8	0.8	0.9
- Relationships between pupils	0.8		0.5
- Support from the teacher and relationship between teacher and pupils	0.8		0.8

3.2.2 *The Cito Pupil Monitoring System*

Schools' average pupil achievement level was measured by means of spelling and mathematics tests from the pupil monitoring system (LVS) developed by Cito (the Dutch Testing and Measurement Institute). Cito developed the LVS to monitor pupils' achievement in primary schools (age 4-12) over time.

The LVS includes three interrelated portfolios for most subject areas and focuses on basic skills. The pupils are tested twice a year: in January and at the end of each school year, in June. The test results may be represented in one measure, and therefore progress of an individual pupil may be followed systematically during his or her primary school career (www.citogroep.nl; Vlug, 1997).

The tests for spelling and mathematics were chosen because they are available for the grades 3 to 8, and pupils are monitored from grades 3/4 to grades 7/8. The tests were administered by the schools.

The test scores for spelling and mathematics were collected before schools used ZEBO for the first time (pre-test) in June 2002. Test scores were again collected for the years 2003 (group 4/5), 2004 (group 5/6), 2005 (group 6/7), and 2006 (group 7/8) (post-tests). However, most schools chose to use the Cito elementary school leavers' attainment tests in grade 8 instead of the LVS tests. Therefore, in 2006, only the grade 7 LVS test scores could be obtained.

Reliability and Validity of the LVS Tests

The quality of (almost) all available Dutch achievement tests is judged by the Committee Test Affairs Netherlands (COTAN) (Evers et al., 2002). Their judgment is based on the following 7 criteria: basic assumption about test construction, quality of the testing material, quality of the instructions, standards, reliability, construct validity and criterion validity. All LVS tests used were assessed as good on the first five criteria. With respect to the sixth criterion, construct validity, the LVS tests used were assessed as adequate. With respect to the criterion validity it was not possible to give a judgment. According to the authors/publisher the tests are not meant to make predictions.

3.2.3 The Pupil Form

At the beginning of the school year 2003/2004 teachers were asked to fill out a pupil form to gather information about the following pupil background characteristics which may influence pupil achievement: gender, socio economic status (SES) (measured with pupil "weight"), language at home (Dutch, Dialect, Turkish and other), perceived intelligence (low, average, or high), class size, and age.

Boys perform better in some subjects, and girls perform better in others. Therefore information on gender of each pupil participating was collected.

Data on SES of the pupils were collected by means of the "weight" of pupils. Each pupil in primary education in the Netherlands receives a certain "weight", based on the educational level of the parents and the land of birth of the parents. A native Dutch pupil with poorly educated parents counts for 1.25 pupil weight, an ethnic minority pupil with poorly educated parents counts for 1.9 weight, bargees' children and gipsy children count for respectively 1.4 and 1.7 weight. The weight these pupils receive determines the amount of extra money a school receives for staffing (www.minocw.nl).

Findings from a study into primary schools indicate that younger members of a year group attain lower achievement scores (Mortimore, 1998), so birth-dates of the pupils were also collected.

Another pupil background measure is the home language of the pupils. These data were collected because home language may influence the achievement scores on the spelling test. Finally, teachers were asked to give an estimate of the IQ of each pupil in their class, in comparison with the national mean. To avoid increasing the workload of schools by conducting standardised IQ tests, teachers were asked to assess the intelligence of each pupil in their class. With respect to the reliability of the use of teacher judgments about the intelligence of their pupils, research (Follman, 1991; Wild, 1993; Tellegen, Winkel, Wijnberg-Williams & Laros, 1998; Biesheuvel & Flim, 2001) shows that teachers are more or less able to judge the intelligence of their pupils, but their judgment cannot be used as a substitute for intelligence measured by an intelligence test. Folman (1991), for example, investigated the correlations between teachers' estimates and pupils' standardised IQs across 32 articles. The correlations presented in those articles ranged from 0.25 to 0.88 with a median of 0.55, a considerable congruence. However, the minimum for such a correlation to be sufficient is 0.80. Despite this, since administration of standardised IQ tests on all 3.220 pupils participating in this study was impossible, it was decided to use teachers' judgments of the intelligence of their pupils. In this study the term "perceived intelligence" is used to indicate that intelligence was estimated by teachers' judgment.

3.2.4 Evaluation of ZEBO Questionnaire

In order to study how schools use ZEBO output, what the effects (other than the effect on pupil achievement) of ZEBO are, and which factors influence the use of ZEBO, the Evaluation of ZEBO Questionnaire) was developed. Since the research group is quite large (in 2002, for example, 220 teachers and 41 principals from 50 schools completed the questionnaire) a questionnaire was considered ideal as this facilitated the study of many variables.

To answer the research questions in this study, a new framework, based on Visscher's (2002) framework, was developed. The Evaluation of ZEBO Questionnaire was devised, based on this new framework. All variables discussed in section 2.2 to section 2.4 were included in the questionnaire. The questionnaire may be found in Appendix 3.2.

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 (see Appendix 3.2).

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

- § Characteristics of ZEBO: this scale is comprised of 10 items, assessing for example, the perceived clarity and relevance of the output [items A1-A10];
- § Implementation process features: 8 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 [items B1-B7];
- § School organisational characteristics: 20 items were formulated for this scale, measuring, for example, the innovation attitude of staff and time and resources available for innovation activities [items C1-C3];

- § ZEBO use: this scale consists of 21 items evaluating, for example, discussion of the output and measures taken as a result [items D1-D7];
- § The effects of ZEBO use: This scale includes 10 items, appraising such effects of ZEBO use as those on pupil care and adaptive education, for example [items E2.1-E3].

For almost all questionnaire items, a statement format with a four 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 (with the exception of items C1a, E1 and E2). The direction of some items was reversed to prevent response bias.

Reliability and Validity of the Evaluation of ZEBO Questionnaire

Factor and reliability analyses revealed eight scales in the questionnaire (Table 3.5). The internal consistency (Cronbach’s alpha) of the questionnaire was assessed, which showed that most scales are sufficiently reliable. However, the initially developed scales for ‘innovation capacity’ and ‘innovation attitude’ were not sufficiently reliable. Factor analyses showed that several factors of the initial ‘innovation capacity’ scale fit more satisfactorily in the ‘innovation attitude’ scale. However, the ‘innovation attitude’ scale at principal level was still not sufficiently reliable, when applying the general rule that the reliability of a scale is ‘good’ if $\alpha \geq 0.8$ and ‘sufficient’ if $0.6 \leq \alpha < 0.8$ (De Heus, Van Der Leeden & Gazendam, 1995). This scale could not be improved by removing items. These scales were sufficiently reliable at teacher level. The low alpha may also have been caused by the fact that only 41 principals completed the questionnaire. The ‘innovation attitude’ scale was not reliable at the aggregated school level, and the ‘pressure and promoting implementation factors’ scale was not reliable at teacher level. Information on the reliability of these scales is presented in Table 3.5.

Reliability increased in the questionnaire of 2004. More respondents filled out the questionnaire at that time and this probably increased the reliability. The ‘innovation attitude’ scale, at principal level, and ‘pressure and promoting factors’ scale, at teacher level, however, were still found to be insufficiently reliable. Reliability remained, with the afore mentioned exceptions, sufficient to good. The only problem scale throughout, was the ‘pressure and promoting factors’ scale, at teacher level (see Appendix 3.3). However, the only scale scores which were actually used in the analyses, were the ZEBO use scales. As well as these use scales, individual items from the questionnaire were used in the analyses. The “relevance of output” and “clarity of output” variables were, for example, used in the analyses instead of the general “characteristics of ZEBO” scale.

To ensure the face validity of the instrument the questionnaire was piloted on a small scale. A teacher and a principal from one school, within the research group, went through the questionnaire and gave comments on complexly phrased questions and those which were not clear. Another teacher who was not familiar with ZEBO also commented on the questionnaire. Based on the comments of these three people a few adjustments were made (rephrasing some questions).

Table 3.5 Reliability of the scales in the questionnaire in 2003

Reliability at	Principal level: Cronbach's α (items) 2003	Teacher level: Cronbach's α (items) 2003	Aggregated school level: Cronbach's α (items) 2003
Scale	(N=41)	(N=220)	(N=50)
Characteristics of ZEBO	0.6 (10)	0.8 (10)	0.8 (10)
Implementation process features: training and support	0.6 (4)	0.9 (4)	0.8 (4)
Implementation process features: Pressure and promoting factors	0.7 (3)	0.5 (3)	0.7 (3)
School organisational features: innovation attitude	0.3 (7)	0.6 (7)	0.5 (7)
School organisational features: innovation capacity	0.7 (12)	0.8 (12)	0.9 (12)
ZEBO use	0.8 (9)	0.9 (9)	0.9 (9)
Conceptual use of ZEBO	0.7 (4)	0.8 (4)	0.9 (4)
Instrumental use of ZEBO	0.6 (5)	0.7 (5)	0.8 (5)
Effects of the use of ZEBO	0.9 (9)	0.9 (9)	0.9 (9)

Response on the First Evaluation (2003)

The Evaluation of ZEBO Questionnaire was sent to the project schools for the first time in May 2003, three months to almost a year after schools had used ZEBO for the first time (the former are schools which were not able to use ZEBO earlier because of technical problems). All 64 schools which worked with ZEBO received the questionnaire for the teachers (grade 1 to 8) and principals. Fifty schools (78%) returned (part of) the nine questionnaires which were sent. Four schools sent back only one questionnaire, and four schools returned all nine questionnaires. On average, five questionnaires per school were sent back. From the 576 questionnaires which were sent, 261 were returned (45.3%).

In Appendix 3.4, an overview of the functions of the respondents who sent back the questionnaire in each phase of the study and the number of questionnaires which were returned may be found. Some schools did not return any of the questionnaires for several reasons:

- § Merger with other schools;
- § Relocation;
- § The ZEBO output was lost;
- § Staff changes (some of the respondents who had worked with ZEBO were not working in the school anymore);
- § Illness of respondents during the ZEBO administration.

Furthermore, during the first ZEBO administration only the principal and teachers in grades three and four from several schools had worked with ZEBO (as was requested by the project, since the first administration was planned for June, which is a very busy period for schools) and consequently only these three respondents could complete the questionnaire.

Response on the Second Evaluation (2004)

The Evaluation of ZEBO Questionnaire was sent to the 58 schools two months after schools administered ZEBO. Most schools used ZEBO at the beginning of 2004, others later during that year (in September at the latest). Fifty schools (86%) returned some of the nine questionnaires which were sent to them. From the 522 questionnaires which were sent to schools, 284 were returned (54.4%). Three schools sent back only one questionnaire, and eight schools sent back all nine questionnaires. On average, six questionnaires per school were sent back. The reasons given for not returning the questionnaires were the same as those given in 2003.

Response on the Third Evaluation (2006)

The Evaluation of ZEBO Questionnaire was sent to the 43 schools which had used ZEBO, two months after they had completed the administration of ZEBO. Most schools used ZEBO at the end of 2005, others at the beginning of 2006. Thirty-one (72.1%) returned some of the nine questionnaires which were sent to them. Of 387 questionnaires which were sent, 166 were returned (42.9%). Two schools sent back all the questionnaires and four schools sent only one questionnaire back. On average 5 questionnaires per school were returned. Schools mentioned the same reasons for not returning the questionnaires as they had in 2003.

3.2.5 Interviews

Interviews were used to help interpret and explain the quantitative findings of the Evaluation of ZEBO Questionnaire (Steckler et al., 1992), and to enhance internal validity by investigating whether the analyses of the data from the questionnaire and interviews led to the same results (Meijer, Verloop & Beijaard, 2001).

The Interview Schedule and Focus Group Questions

An interview schedule was constructed to gain more insight into the use of ZEBO and to find out whether the interviews led to the same results as the questionnaires. The interview questions were formulated on the basis of the results of the questionnaire. To gain more insight into the nature of the factors influencing the use of ZEBO, respondents were asked in the interviews about the characteristics of ZEBO, the implementation process features, and the school organisational characteristics. Interview questions on the use of ZEBO were also formulated, for example, about the measures which were taken based on the ZEBO output.

The interview questions were tested and adapted slightly for clarity, with the help of a teacher who was not involved in the ZEBO-project, but familiar with it. The interview schedule may be found in Appendix 3.5, including the extra questions which were added to the interview schedule in 2005, to find out whether schools used ZEBO differently the second time.

In 2006, due to time restrictions, it was not possible to conduct interviews with respondents from different schools separately. As an alternative a focus group interview was conducted. Focus groups may be defined as a research technique which collects data through group interaction on a topic determined by the researcher (Morgan, 1996). Focus groups are a form of group interview with a semi-structured questioning approach which relies on participants' response (Litosseliti, 2003). An advantage is that the researcher may get reactions from a relatively wide range of participants in a relatively short time (Morgan, 1996).

According to Morgan (1996) what makes the discussion in focus groups more than the sum of individual interviews is the fact that the participants both query each other and explain themselves to each other. A focus group presents a natural environment in that participants influence, and are influenced by, others, just as they are in real life. Interaction plays an important role in focus groups (Litosseliti, 2003). Focus groups may produce in-depth information on the topic at hand (in this study the use of ZEBO).

Focus groups may be used as a supplementary source of data (Litosseliti, 2003) to check the conclusions from other analyses (Morgan, 1996). The focus group in this study was conducted with that objective in mind: to further validate and complement the conclusions of this study, especially concerning the factors influencing the use of ZEBO. The questionnaires, interviews, and pupil achievement tests provided a clear depiction of the use and effects of ZEBO, but not all the variance in the use of ZEBO could be explained by the factors influencing the use of school self-evaluation results in the framework.

A topic guide was developed to generate a broad yet focussed in-depth discussion on the use of ZEBO. It consists of a list of issues and statements to be explored during the session. The topic guide was based on the conclusions of this study and may be found in Appendix 3.6.

Selection of Schools for Interviews, after the First Time Schools Used ZEBO (2003)

Based on the results of the "use-variables" of the Evaluation of ZEBO Questionnaire in 2003, the 50 schools which worked with ZEBO were divided into three groups of ZEBO users:

- § One group of schools which did not use the ZEBO output (9 schools);
- § Another group of schools which made average use of the results (the results were discussed but few measures were taken; 33 schools);
- § A third group of schools which used the ZEBO output intensively (measures were taken based on the ZEBO output to improve the quality of education; 8 schools).

The items E1a to E1g (see Appendix 3.2) were re-scaled into one variable on a four-point response scale. The same was done for items E2a to E2g. A new variable, "total use", was created by computing the closed use items including the two re-scaled items (a total of 9 variables). Schools which scored between 9.00 (the minimum) and 11.3 on

this variable (mean minus one standard deviation) received the label “Low Self-Evaluation (LoSE)”. Schools which scored between 11.4 and 23.1 (mean plus one standard deviation) received the label “Average Self-Evaluation (AvSE)”, and schools which scored between 23.2 and 36.0 (the maximum) received the label “High Self-Evaluation (HiSE)” (see Table 3.6).

Ten principals and twenty-one teachers (from 11 schools) were interviewed to further explore the differences between the LoSE, AvSE and HiSE schools. From the HiSE schools category only two schools were willing to participate. Interviews were conducted with respondents of five schools in the category AvSE schools and with respondents from four schools in the category LoSE schools (more details may be found in Appendix 3.7). The interviews with the principal and the teachers of the followed pupil cohort took place in November/December 2003. All interviews were taped and notes were taken, both with the permission of participants. Interviews lasted from 30 minutes to one hour.

Selection of Schools for Interviews, after the Second Time Schools Used ZEBO (2004)

Nine schools were selected for interviews on the basis of the Evaluation of ZEBO Questionnaire in 2004. The selection was made in the same way as in 2003. Schools which scored between 9.00 (the minimum) and 12.4 on the newly constructed use variable (mean minus one standard deviation) received the label LoSE. Schools which scored between 12.5 and 23.6 (mean plus one standard deviation) received the label AvSE, and schools which scored between 23.7 and 36.0 (the maximum) received the label HiSE (see Table 3.6).

From each category of schools, three were selected for interviews. Of the fifty schools which returned the questionnaire, 7 were identified as LoSE, 8 schools as HiSE and the remaining 35 were labelled as AvSE. Seven principals and 18 teachers were interviewed (for more details see Appendix 3.7) in June/July 2005.

Selection of Schools for a Focus Group after the Third Time Schools Used ZEBO (2006)

To select respondents for the focus group again schools were divided in three groups of ZEBO users based on the questionnaire results. Schools which scored between 9.0 and 14.1 received the label “LoSE” (6 schools). Those which scored between 14.2 and 22.9 (21 schools) received the label “AvSE”. Schools which scored between 23.0 and 36.0 received the label “HiSE” (see Table 3.6).

From the LoSE schools, four participants from four schools intended to participate in the focus group. Three respondents from three AvSE schools and three participants from three HiSE schools intended to participate in the focus group, making a total of ten planned participants (see Appendix 3.7 for more details on the participating respondents). However, in November when the focus group met, only two participants were present. The other eight participants could not attend, due to sickness and other pressing issues. Only two principals (from a HiSE school and an AvSE school) were

interviewed (simultaneously). The interview questions focussed on the question why, according to these respondents, several schools did not use the ZEBO output to improve the quality of education.

Table 3.6 *Three groups of ZEBO users*

	LoSE	Nr. of schools	AvSE	Nr. of schools	HiSE	Nr. of schools
2003	9.0 – 11.3	9	11.4 – 23.1	33	23.2 – 36.0	8
2004	9.0 – 12.4	7	12.5 – 23.6	35	23.7 – 36.0	8
2006	9.0 – 14.1	6	14.2 – 22.9	21	23.0 – 36.0	4

3.2.6 *Mixed Methods*

In this study, mixed methods were used: both quantitative data and qualitative data were employed. Burke Johnson and Onwuegbuzie (2004) define mixed methods research as the class of research in which the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language in a single study. An advantage of using mixed methods is that if findings are corroborated across different approaches, then greater confidence may be held in the conclusion. This is called triangulation. A finding may be verified by showing that independent measures of it agree with it, or, at least, do not contradict it (Miles & Huberman, 1994; Burke Johnson & Onwuegbuzie, 2004; Mertens, 1998).

One reason for not using quantitative research alone, is that the categories and theories used by the researcher, may not reflect the understanding of the local constituencies. The researcher may also inadvertently, because of the focus on theory and hypothesis testing, miss out on some phenomena. Some of the advantages of qualitative research are that the data are based on the participants' own categories of meaning, the data provide understanding and description of people's personal experience of phenomena, and qualitative research may be used for studying a limited number of cases in depth (Burke Johnson & Onwuegbuzie, 2004).

Another advantage of using mixed methods is complementarity. Complementarity refers to seeking elaboration, enhancement, illustration and clarification of the results of one method with results of the other method (Meijer et al., 2001). The data from the interview results combined with the data from the questionnaires formed a more comprehensive view of the phenomenon studied, the use of ZEBO. In Appendix 3.8 an overview of all the factors studied and the instruments used may be found.

Table 3.7 presents an overview of the data collection, including details on the time and the level at which data were collected along with the goals of data collection. In Appendix 3.9 a timeline of the research is presented, with an overview of all the data collected.

Table 3.7 Overview of the goals, the level, and the timeframe of the data-collection

Goal	Data collected at pupil level	Data collected at teacher level	Data collected at principal level
Explore the use of ZEB0		Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006	Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006
		Interviews conducted in 2003 and 2005	Interviews conducted in 2003 and 2005
		Focus group conducted in 2006	Focus group conducted in 2006
Study the effects of ZEB0 on pupil achievement	Pupil background characteristics in 2002	Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006	Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006
	Pupil achievement spelling and mathematics pre-tests in 2002		
	Pupil achievement spelling and mathematics post-tests in 2003, 2004, 2005, and 2006		
Investigate the (un)intended effects of the use of ZEB0		Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006	Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006
	Assess the variables influencing the use of ZEB0	The results of the ZEB0 administration in 2002/2003, 2004, and 2005	The results of the ZEB0 administration in 2002/2003, 2004, and 2005
		Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006	Evaluation of ZEB0 Questionnaire in 2003, 2004, and 2006
		Interviews conducted in 2003 and 2005	Interviews conducted in 2003 and 2005
			Focus group conducted in 2006

3.3 Data-Analyses

How and to What Extent do Schools Use ZEB0?

In order to answering this question, a distinction was made between instrumental use and conceptual use of ZEB0. A “conceptual ZEB0 use” scale was constructed by calculating the sum of the following elements (from the Evaluation of ZEB0 Questionnaire, all with a four point response scale):

- § The degree to which the use of ZEB0 provided the user with new insights;
- § The degree to which the ZEB0 output highlighted problems within the school;
- § The degree to which school staff devised solutions for the problems highlighted by ZEB0;
- § The degree to the respondent devised solutions for the problems highlighted by ZEB0.

An “instrumental use of ZEB0” scale was constructed by calculating the sum of the following aspects:

- § The degree to which school staff studied the ZEB0 output;
- § The degree to which measures were taken on the basis of the various ZEB0 outputs;

- § The degree to which the ZEB0 output was discussed within the school;
- § The degree to which, on the basis of ZEB0 output, school staff took measures to improve the quality of education;
- § The degree to which, on the basis of ZEB0 output, the respondent took measures to improve the quality of education.

The Conceptual use of ZEB0 scale includes four variables and has a minimum score of 4 and a maximum of 16. Schools which scored 10 or higher on this scale were considered to use the results conceptually. Respondents from these schools answered on all four questions concerning the conceptual use of ZEB0 on average “to a moderate degree”, “or to a great degree”. The Instrumental use of ZEB0 scale includes 5 variables and has a minimum score of 5 and a maximum of 20. Schools which scored 12.5 or higher on this scale were considered to use the results instrumentally. Respondents from these schools answered on average on all five instrumental use variables “to a moderate degree” or “to a great degree”.

Furthermore, the data obtained by means of the individual and focus group interviews were analysed according to the qualitative analysis method of Baarda, De Goede, and Teunissen (2000). Their analysis method consists of a number of steps which were followed:

- § Segmenting the data: each fragment handles one subject and may be read in the context of the interview text;
- § Coding the segments: each fragment received a code relevant for its content and for the research objectives, each code includes all relevant aspects of that fragment, and each code tells something about the individual or the situation;
- § Removing irrelevant segments: data which did not contain relevant information for answering the research questions were removed;
- § Finally, the codes and their relationships were described to answer the research questions.

The program Atlas/ti was used to code the interview transcripts and focus group notes. All codes used, reflect variables from the theoretical framework. After coding (and recoding), segments were sorted into meaningful subsets, called “families”, to make it easier to deal with the large number of codes. Five families were formed, based on the theoretical framework: characteristics of ZEB0, as perceived by its users, implementation process features, school organisational characteristics, ZEB0 use, and the effects of ZEB0 use. Categories and codes within these families were used to analyse the interviews in detail. An overview of the families created and the categories with the number of codes which were used may be found in Appendix 3.10. The next step was to create networks in which the families, factors and codes were connected. First, a network was created for each separate school in order to study the use of ZEB0. Next, three networks were created: one network for the schools which did not use the ZEB0 output, one network for the schools which made average use of ZEB0, and one network for schools which made intensive use of ZEB0. To answer the research

question about which factors influence the use of ZEB0, these networks were compared to each other. Figure 3.1 shows an example of a network.

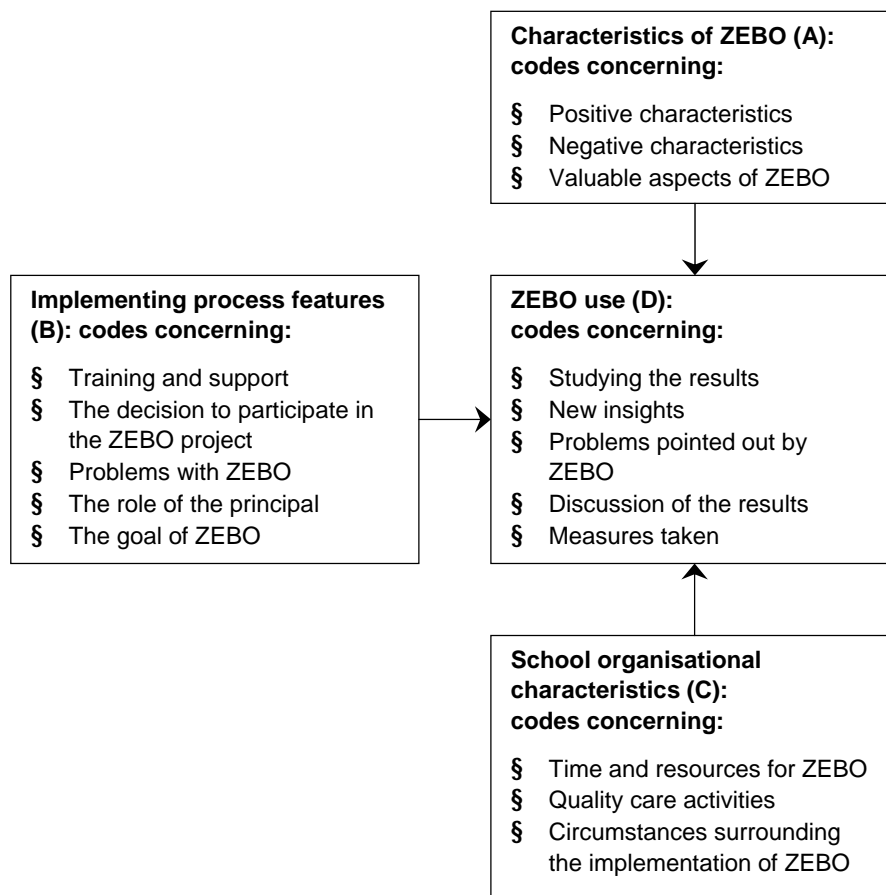


Figure 3.1 Example of a network

To establish reliability, interrater-agreement was calculated for each of the codes used. A second independent coder coded 240 of the 945 interview fragments (25%). The categories were given to the second coder, since the interview questions already revealed these. For (the) coding (of) the different interview fragments Cohen’s Kappa yielded between 0.63 and 1.00. An overview of Cohen’s Kappa for all the categories may be found in Appendix 3.11. An interrater agreement of 0.60 or higher is considered acceptable or substantial, and an interrater agreement of 0.80 or higher is considered to be “good” or almost perfect (Eggen & Sanders, 1993).

What Are the Effects of the Use of ZEB0 on Pupil Achievement?

As the pupil achievement data have a nested structure (pupils are nested within classes, and teachers are nested within schools) multilevel analyses were conducted with the help of the MLWin software package. The influence of the conceptual and instrumental use of ZEB0 on spelling and mathematics achievement was analysed per grade, three times, each year after schools had used ZEB0.

After the first evaluation, in the multilevel analyses two models were compared to each other: a basic model, which includes all significant covariates at pupil level and a model which is the same as the basic model, with the exception that a variable for ZEBO use was added. By means of the multilevel analyses, whether or not the model with the added ZEBO-use variable fits the data significantly better than the basic model, was studied. Alpha was set at 0.05 for a 2-tailed test. Variables were left in the model if there was a significant ($p < 0.05$) reduction in the likelihood-ratio statistics (LRS) using chi-squared tables.

After the second and third evaluations multilevel analyses with repeated data were conducted, since the pupils had been measured on more than one occasion. For the first evaluation, data from two years were available. Data from two years are not enough to perform multilevel analyses with repeated measures, since trends in performance over time cannot be studied with data from only two years (Malin & Linnakylä, 2001). Therefore after the first evaluation, a two-level model was used instead of a three-level model. In this model, the pupils constituted the level 1 units and the classes are the level 2 units. For the second and third evaluation, data from three or more years were available. Therefore multilevel analyses with repeated data were conducted; the repeated tests being the level 1 units, the pupils being the level 2 units, and the classes being the level 3 units. Schools were supposed to administer ZEBO three times¹. Therefore the ultimate question was whether the aggregate of the use of ZEBO over the years was related to pupil progress over the years. To answer this question, multilevel analyses with repeated measures were conducted with aggregated ZEBO use variables. These multilevel analyses were first conducted with *spelling* achievement as the dependent variable. Next, analyses were conducted with *mathematics* achievement as the dependent variable. In both analyses, the aggregate conceptual ZEBO use was entered as an independent variable. The aggregate conceptual use variable was constructed by taking the conceptual use of ZEBO in 2003, 2004, and 2006 together. This aggregate variable, calculated for each school, was entered into the multilevel analyses as a continuous variable. Analyses with instrumental ZEBO use as the independent variable were also conducted with spelling achievement as the dependent variable, and mathematics achievement as the dependent variable. By adding the sum of instrumental use of ZEBO per school in 2003, 2004, and 2006, an aggregate instrumental use variable was calculated.

What Are the Other Effects of ZEBO Use?

Part of the data, obtained by means of the Evaluation of ZEBO Questionnaire, was used to determine the intended (such as an improvement in educational leadership) and unintended effects (such as stress) of ZEBO use. To study the effects of the use of ZEBO, the data for (un)intended effects was analysed for each moment of measurement (T1, T2 and T3) by computing frequencies.

¹ However, not all schools administered ZEBO three times. Thirty-six schools administered ZEBO three times, twenty-two schools two times, and 10 schools administered ZEBO only one time.

Which Factors Influence the Use of ZEBO?

To answer the question about which variables have a relationship with the degree of ZEBO use by principals, correlations were calculated between ZEBO use (D) and, the characteristics of ZEBO (A), implementation process features (B), and school organisational characteristics (C), as measured by means of the Evaluation of ZEBO Questionnaire.

Variables that correlate significantly with the use of ZEBO were thereafter entered as potential predictors into stepwise regression analyses (in order of size) on ZEBO use for the data from the principals. To draw conclusions based on a regression analysis, several assumptions must be met (Field, 2000): the predictors should have some variation in value (e.g. non-zero variance), there should be no perfect linear relationship between two or more of the predictors (e.g. no perfect multicollinearity), at each level of the predictor variables, the variance of the residual terms should be constant (e.g. homoscedasticity), for any two observations the residual terms should be uncorrelated (e.g. independent errors), the residuals in the model should be random, normally distributed variables with a mean of zero, the values of the outcome variable should be independent, and the mean values of the outcome variable for each increment of the predictors lie along a straight line. Some variables did not meet these assumptions and were therefore removed from the analyses. For the remaining variables and scales used, the assumptions were met.

The number of available cases (N = 41 in 2003, N=48 in 2004. and N=25 in 2006) required a selection of potential predictors. A maximum of 4 variables (2 in 2006) was entered in each regression analysis: those variables which proved to correlate most strongly and significantly with the use of ZEBO. We do recognize the fact that this may result in chance capitalisation. MacCallum, Roznowski and Necowitz (1992) state, in this regard, that data-driven search processes are susceptible to capitalisation of change in that characteristics of the sample may influence the analyses conducted. When a sequential specification search is conducted using data from a single sample, researchers may not be sure that the specific model generalises beyond that sample. The model may be a good description of the data set used, but it may not be a good description of any other data set. Furthermore, a variable may not have a bivariate correlation with the dependent variable, whereas, if another variable is partialled out, it may have.

Considering the number of variables in the theoretical framework, which are all expected to influence the use of ZEBO, it was necessary to come to a reduction of variables (to be entered into the regression analyses). Since it was not possible to come to a theoretical reduction because all variables seemed equally important, it was decided to make a reduction based on the correlational analyses.

Multilevel analyses instead of regression analyses were conducted on the data from the teachers, to test the hypothesized relationships in the theoretical framework. Because the teacher data collected in this study have a nested structure (teachers are nested

within schools) multilevel analysis was required for the data from teachers.

It was decided not to compute multiple regression or multilevel analyses at the level of the school as a whole (taking together the data based on the principals and teachers data), since variance analyses showed that teachers and principals differ significantly in their opinion on the characteristics of ZEBO, the implementation process features, and the characteristics of the school organisation. Aggregating the data to the level of the school would mean a considerable loss of information. According to Snijders and Bosker (1999) aggregation may also result in gross errors.

Finally, the data obtained in the interviews were analysed according to the qualitative analysis method of Baarda, De Goede and Teunissen (2000) as described above to further explore the factors influencing the use of ZEBO. Chapters 4, 5 and 6 present the results of the analyses described above.

Chapter 4

The Use of ZEBO

4.1 Introduction

The first goal of this study was to systematically acquire detailed knowledge on the use of ZEBO output within schools. Figure 4.1 shows the variables studied in order to assess the use of ZEBO output. Following Weiss (1998a), a distinction was made between conceptual use and instrumental use of evaluation findings. Rossi et al. (1999) define instrumental use as the direct use of evaluation findings: the results are analysed and actions are based on the interpretation of these results. The use of evaluation findings may also be conceptual: the evaluation feedback influences thinking about issues. This distinction was used in studying how schools used ZEBO.

ZEBO use	
Conceptual use:	
§	The degree to which the ZEBO output provides the user with new insights;
§	The degree to which the ZEBO output highlights certain problems;
§	The degree to which school staff devise solutions for the problems highlighted by ZEBO;
§	The degree to which the respondent devises solutions for the problems highlighted by ZEBO.
Instrumental use	
§	The degree to which the various types of ZEBO output are studied;
§	The degree to which measures are taken based on various types of ZEBO output;
§	The degree to which the ZEBO output is discussed within the school;
§	The degree to which school staff takes measures to improve the quality of education based on the ZEBO output;
§	The degree to which the respondent takes measures to improve the quality of education.

Figure 4.1 The variables studied in order to assess the use of ZEBO output

To study how schools use ZEBO, a questionnaire was developed (“the Evaluation of ZEBO Questionnaire”). Principals and teachers completed this questionnaire in 2003, 2004, and 2006, each time after schools had administered ZEBO. The results obtained from the Evaluation of ZEBO Questionnaire are presented in section 4.2.

An interview schedule was constructed to gain deeper insights into the use of ZEBO and to validate the outcomes of the Evaluation of ZEBO Questionnaire. The interview questions were based on the results of the questionnaire. The results of the interviews are discussed in section 4.3. In section 4.4, the evaluation results for 2003, 2004, and 2006 are compared with each other, and the chapter ends with a summary and discussion of the results of ZEBO use (section 4.5).

4.2 ZEBO Use: The Results from the Evaluation of ZEBO Questionnaire

In 2003, 64 schools administered ZEBO. Respondents from 50 schools completed the Evaluation of ZEBO Questionnaire that year. In 2004, 58 schools administered ZEBO, and 48 principals and 236 teachers completed the questionnaire at that time. Fewer schools (43) schools administered ZEBO in 2006¹. Respondents from 31 schools filled out the questionnaire (see Table 4.1).

Table 4.1 *ZEBO administration and the number of evaluation of ZEBO questionnaires returned*

Year	ZEBO administration in N schools	Questionnaires from N schools	Questionnaires from N principals	Questionnaires from N teachers
2003	64	50	41	220
2004	58	50	48	236
2006	43	31	25	141

The questionnaire results are presented below. The results of the first evaluation of ZEBO use (2003) are described in detail. For the second (2004) and third (2006) evaluations a summary of the results is presented, focusing on the differences between ZEBO use during the three years.

4.2.1 Conceptual Use of ZEBO output in 2003

Thirty-four teachers and principals (13%, N=261) from six schools (12% of the schools) reported that they used ZEBO conceptually: the results influenced their thinking. They stated that ZEBO output increased their awareness of the weaker and stronger aspects of their functioning. These respondents also indicated that the ZEBO output led to new insights. The ZEBO output, for example, made teachers more aware of the way they are perceived by their pupils. The ZEBO output also highlighted specific problems and the respondents devised solutions for the problems which were found.

The results of the conceptual use of ZEBO are presented in Table 4.2 (percentages are shown in round numbers). A distinction was made between the results from the principals and the results from the teachers. Principals made more conceptual use of ZEBO output than teachers, but in general the conceptual use of ZEBO output was limited. Nineteen principals (47%) indicated that the results from ZEBO led, to a great degree or to a moderate degree, to new insights. Principals mentioned, for example, that ZEBO output highlighted how teachers judge the school and the principal. Twenty-nine teachers (13%) indicated that the results led to new insights, for example, insight into how the pupils judge classroom instruction. Thirteen principals (32%) indicated that, to a moderate degree, the results highlighted certain problems. Sometimes both principals and the school team devised solutions for these problems. Teachers also indicated that ZEBO highlighted certain problems (22%, n=220) and that the school team devised solutions for these problems (19%, n=220).

¹ This decrease in ZEBO administration had two main reasons: nine schools stopped participating in the project and twelve schools chose to postpone administrating ZEBO until 2007. These twelve schools have included ZEBO in their quality care cycle.

Table 4.2 *Conceptual ZEBO use by principals and teachers*

Questions	Respondents	N	Year	To a great degree		To a moderate degree		To a small degree		To a minimal degree/not		Missing/ I do not know	
				%	(n)	%	(n)	%	(n)	%	(n)	%	(n)
ZEBO use provided me with new insights	Principals	41	2003	5	(2)	42	(17)	32	(13)	20	(8)	2	(1)
		48	2004	4	(2)	35	(17)	23	(11)	33	(16)	2	(1)
		25	2006	9	(2)	48	(12)	20	(5)	24	(6)	0	(0)
	Teachers	220	2003	2	(5)	11	(24)	14	(31)	53	(116)	20	(44)
		236	2004	3	(8)	15	(35)	25	(58)	47	(111)	10	(24)
		141	2006	6	9	28	(40)	28	(39)	31	(43)	7	(10)
The ZEBO output highlighted certain problems	Principals	41	2003	2	(1)	32	(15)	42	(17)	24	(10)	0	(0)
		48	2004	2	(1)	27	(13)	35	(17)	33	(16)	2	(1)
		25	2006	4	(1)	36	(9)	32	(8)	24	(6)	4	(1)
	Teachers	220	2003	6	(13)	16	(34)	24	(52)	28	(61)	27	(60)
		236	2004	0	(1)	17	(39)	25	(59)	33	(78)	25	(59)
		141	2006	4	(6)	16	(23)	26	(37)	38	(54)	15	(21)
I devised Solutions for these problems	Principals	41	2003	2	(1)	34	(14)	29	(12)	34	(14)	0	(0)
		48	2004	0	(0)	21	(10)	31	(15)	42	(20)	6	(3)
		25	2006	8	(2)	24	(6)	20	(5)	44	(11)	4	(1)
	Teachers	220	2003	1	(2)	3	(6)	8	(17)	72	(159)	16	(36)
		236	2004	0	(0)	5	(12)	11	(25)	72	(170)	12	(29)
		141	2006	1	(2)	5	(7)	17	(23)	69	(97)	9	(12)
School staff devised solutions	Principals	41	2003	7	(3)	24	(10)	32	(13)	32	(13)	5	(2)
		48	2004	2	(1)	19	(9)	27	(13)	46	(22)	6	(3)
		25	2006	0	(0)	48	(12)	12	(3)	36	(9)	4	(1)
	Teachers	220	2003	4	(9)	15	(32)	13	(29)	47	(103)	21	(47)
		236	2004	2	(5)	13	(31)	17	(40)	52	(123)	12	(29)
		141	2006	1	(2)	18	(25)	18	(25)	45	(64)	18	(25)

4.2.2 *Conceptual Use of ZEBO output in 2004 and 2006*

In 2004, twenty principals and teachers from four schools (8% of the schools) used ZEBO output conceptually. In 2006, nineteen teachers and principals from three schools (10% of the schools) used ZEBO output conceptually. As in 2003, principals indicated that ZEBO output also led to new insights in 2004 and 2006.

Relatively, the conceptual use of ZEBO increased over the years, but in absolute numbers, the use decreased slightly due to the decrease in the number of schools which had administered ZEBO in 2006.

In 2004 and 2006 more teachers attended to the ZEBO output. In 2003, only 11% of the teachers (24 teachers) had indicated that ZEBO output led to new insight to a moderate degree. This percentage had increased in 2006 to 28% (40 teachers) (see Table 4.2). The new insights related to:

- § Aspects of the school's functioning which needed improvement;
- § The opinions of pupils;
- § Differences of opinions or agreements within the team;
- § Differences on ZEBO scales compared with the national average.

In response to one of the open questions² one of the teachers stated in the Evaluation of ZEBO Questionnaire: “our school is functioning well in comparison with the national mean. The ZEBO output confirms that we are on the right track”. Another teacher stated “it makes you more aware of the weaker aspects of the school’s functioning”. Teachers appreciated the opinion of their pupils. One teacher stated, for example, “I put too much pressure on the children. My expectations are too high”.

ZEBO highlighted certain problems to a great or moderate degree according to 14 principals (29%) and 40 teachers (17%) in 2004. The degree to which the school team devised solutions for these problems decreased slightly in 2004, compared with 2003, but increased again in 2006. In 2006, 12 principals (48%) and 27 teachers (19%) indicated that the school team devised solutions for the problems, to a great or moderate degree.

4.2.3 Instrumental Use of ZEBO output in 2003

Thirty-one teachers and principals (12%, N=261) from twelve schools (24% of the schools) used ZEBO instrumentally (these schools also used ZEBO conceptually). These respondents studied the ZEBO output, discussed the results, and on that basis measures were taken to improve the quality of education, such as:

- § More frequent and open consultation about the quality of education;
- § Stimulating independent pupil learning (for example, by means of block teaching: a method of restructuring lessons to give pupils more time in class);
- § More frequent evaluation of pupil achievement;
- § Implementing classroom consultation;
- § Using the results from ZEBO for the school plan and school prospectus;
- § More explanation for pupils who require this;
- § Accommodating more to differences;
- § More clarity of classroom instruction.

Table 4.3 shows the differences between the results of the principals and the results of the teachers, relating to the instrumental use variables. Instrumental use of ZEBO was limited, especially the instrumental use of ZEBO output by teachers. The majority of principals (81%, n=41) reported that they studied the ZEBO output to a great degree. However, as much as 44% of the teachers (n=220) indicated that they had not studied the ZEBO output at all, or to only a minimal degree. Most principals (64%) stated that the ZEBO output was discussed to a moderate or great degree; however, only slightly over half that percentage of teachers (39%, n=86) agreed. The majority of teachers (51%, n=220) and a substantial minority of principals (32%, n=41) indicated that school staff took no or almost no action based on ZEBO output.

² The questionnaire also included open questions in which respondents were asked to mention examples of, for example, the new insights gained.

Table 4.3 Instrumental ZEB0-use by principals and teachers

Questions	Respondents			To a great degree		To a moderate degree		To a small degree		To a minimal degree/not		Missing/ I do not know	
		N	Year	%	(n)	%	(n)	%	(n)	%	(n)	%	(n)
I studied the various ZEB0 outputs	Principals	41	2003	81	(33)	7	(3)	0	0	12	5	0	(0)
		48	2004	81	(39)	13	(6)	0	0	6	3	0	(0)
		25	2006	68	(17)	12	(3)	8	(2)	12	(3)	0	(0)
	Teachers	220	2003	35	(76)	8	(17)	8	(18)	44	(97)	6	(12)
		236	2004	46	(109)	15	(35)	9	(22)	28	(66)	2	(4)
		141	2006	60	(84)	11	(15)	12	(17)	14	(20)	4	(5)
The ZEB0 outputs were discussed	Principals	41	2003	22	(9)	42	(17)	27	(11)	10	(4)	0	(0)
		48	2004	15	(7)	52	(25)	17	(8)	15	(7)	2	(1)
		25	2006	24	(6)	40	(10)	24	(6)	12	(3)	0	(0)
	Teachers	220	2003	11	(25)	28	(61)	23	(50)	27	(60)	11	(24)
		236	2004	18	(42)	31	(72)	25	(58)	23	(55)	4	(9)
		141	2006	21	(29)	43	(61)	20	(28)	14	(20)	2	(3)
Based on the ZEB0 output I took measures	Principals	41	2003	2	(1)	17	(7)	29	(12)	49	(20)	2	(1)
		48	2004	0	(0)	17	(8)	25	(12)	50	(24)	8	(4)
		25	2006	8	(2)	24	(6)	20	(5)	44	(11)	4	(1)
	Teachers	220	2003	1	(1)	10	(21)	12	(26)	64	(141)	14	(31)
		236	2004	0	(1)	8	(19)	20	(48)	64	(150)	8	(18)
		141	2006	1	(1)	11	(15)	29	(41)	54	(76)	6	(8)
Based on the ZEB0 output school staff took measures	Principals	41	2003	7	(3)	24	(10)	32	(13)	32	(13)	5	(2)
		48	2004	7	(3)	23	(11)	31	(15)	29	(14)	10	(5)
		25	2006	8	(2)	44	(11)	16	(4)	28	(7)	4	(1)
	Teachers	220	2003	5	(10)	16	(34)	10	(22)	51	(112)	19	(42)
		236	2004	3	(7)	13	(31)	17	(41)	53	(125)	14	(32)
		141	2006	4	(6)	20	(28)	23	(33)	42	(59)	11	(15)

Based on the results of the Evaluation of ZEB0 Questionnaire from the first year (2003), the 50 schools which worked with ZEB0 were then divided into three groups³: (for further details see Chapter 3):

- § A group of schools which did not use the output from ZEB0 (Low Self Evaluation, “LoSE”): 9 schools;
- § A group of schools which made average use of the ZEB0 output (Average Self Evaluation, “AvSE”): 33 schools;
- § A group of schools which used the results from ZEB0 intensively (High Self Evaluation, “HiSE”): 8 schools.

³ The ZEB0-using schools were divided into three groups based on the results of all the “use-variables” in the questionnaire. A new variable, “total use” (instrumental and conceptual use together), was created by computing the closed ZEB0 use items. Schools which scored between 9.00 (the minimum) and 11.3 on this variable (mean minus one standard deviation) received the label “Low Self-Evaluation (LoSE)”. Schools which scored between 11.4 and 23.1 (mean plus one standard deviation) were labelled “Average Self-Evaluation (AvSE)”, and schools which scored between 23.2 and 36.0 (the maximum) received the label “High Self-Evaluation (HiSE)”.

Figure 4.2 shows the mean scores of the three groups of ZEBO using schools on elements of ZEBO use (each item is a variable which was measured in the questionnaire). All variables were measured on a four-point scale (1 = to a minimal degree/not, and 4 = to a great degree). Figure 4.2 shows that although the results were discussed, and measures were taken, to some extent, in the HiSE schools, the use of ZEBO output, in general, was limited in 2003.

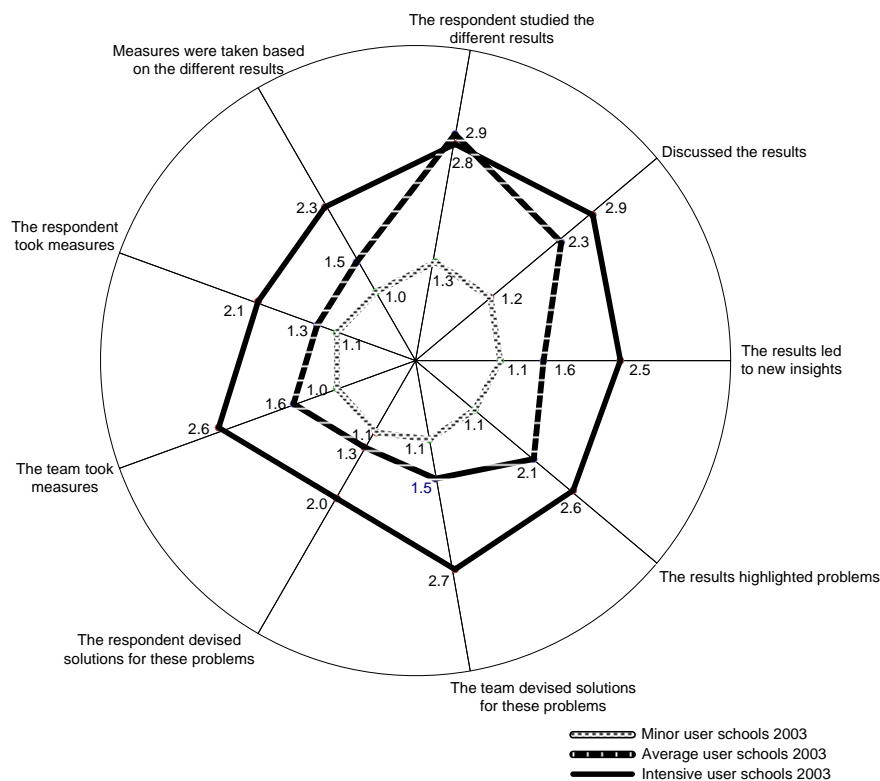


Figure 4.2 Mean scores of the LoSE, AvSE and HiSE schools on each of the ZEBO use variables in 2003 (variables measured on a four-point scale: 1 = not at all, and 4 = to a great degree)

4.2.4 Instrumental Use of ZEBO output in 2004 and 2006

Instrumental use increased relatively and in absolute number in 2004 and 2006, as compared with 2003. In 2004, 43 principals and teachers (15%) from 13 schools (26%) and in 2006, 40 principals and teachers (24%) from 10 schools (32%) used the results instrumentally. Principals studied the results more in 2004, but this decreased slightly in 2006 (see Table 4.3). Yet, as many as 68% of principals (n=17) reported studying the results to a great degree in 2006. The number of teachers who reported studying the ZEBO output increased over the years. In 2003 only 35% of the teachers (n=220) studied the results to a great degree, in 2006 this percentage almost doubled.

The majority of principals (64%, n=41) stated that the results were discussed, to a moderate or great degree, in 2003 as well as in 2004 (67%, n=48) and 2006 (64%, n=25). The degree to which the results were discussed according to teachers increased over the years. In 2003, 39% of teachers (n=220) indicated that the results were discussed to a moderate or a great degree, whereas, in 2006, that percentage was 64% (n=141).

In 2006 a higher number of measures were taken, based on ZEBO output relative to the number taken in 2003 or 2004. Thirteen principals (52%, n=25) and thirty-four teachers (24%, n=141) indicated that the team took measures to improve the quality of education based on the ZEBO output (see Table 4.3). As in 2003, principals primarily mentioned measures at the level of the school organisation, such as more cooperation with teachers, and an increased attention to the professional development of teachers. Teachers primarily referred to measures taken at the level of the classroom, such as differentiation in pupil learning pace and subject matter, and providing pupils with more clarity on classroom rules. Measures which were mentioned in 2004 and/or 2006, but not in 2003 included:

- § Using the help of parents in computer lessons;
- § Stating the objectives of a lesson clearly at the start of each lesson;
- § Evaluating lessons;
- § Paying more attention to gifted pupils;
- § Developing a communication plan;
- § The implementation of new programs, such as BOOM (an adaptive education program).

Based on the Evaluation of ZEBO Questionnaire results in 2004⁴ and 2006⁵, the ZEBO-using schools were again divided into three groups (Table 4.4).

Table 4.4 Number of LoSE schools, AvSE schools and HiSE schools in 2004 and 2006

	LoSE schools	AvSE schools	HiSE schools
2004	7 schools	35 schools	8 schools
2006	6 schools	21 schools	4 schools

Figure 4.3 shows the mean scores of the three groups of schools (LoSE, AvSE, and HiSE) on the various aspects of ZEBO use in 2004. Although the degree to which the ZEBO output was studied increased in 2004 as compared with 2003, the degree to which measures were taken by the teams slightly decreased in 2004, according to respondents from the AvSE schools and HiSE schools.

Figure 4.4 shows the mean scores of the LoSE, AvSE, and HiSE schools on the ZEBO use variables in 2006. ZEBO use increased in 2006 when compared with 2003 and 2004. The ZEBO output was studied and discussed more in schools. Furthermore, the output led to more new insights, and as a result more measures were taken.

⁴ Schools which scored between 9.00 (the minimum) and 12.4 on this variable (mean minus one standard deviation) received the label “LoSE”. Schools which scored between 12.5 and 23.6 (mean plus one standard deviation) were labelled “AvSE”, and schools which scored between 23.7 and 36.0 (the maximum) received the label “HiSE” in 2004.

⁵ Schools which scored between 9.0 and 14.1 received the label “LoSE” (6 schools). Schools which scored between 14.2 and 22.9 (21 schools) were labelled “AvSE”. Schools which scored between 23.0 and 36.0 received the label “HiSE” in 2006.

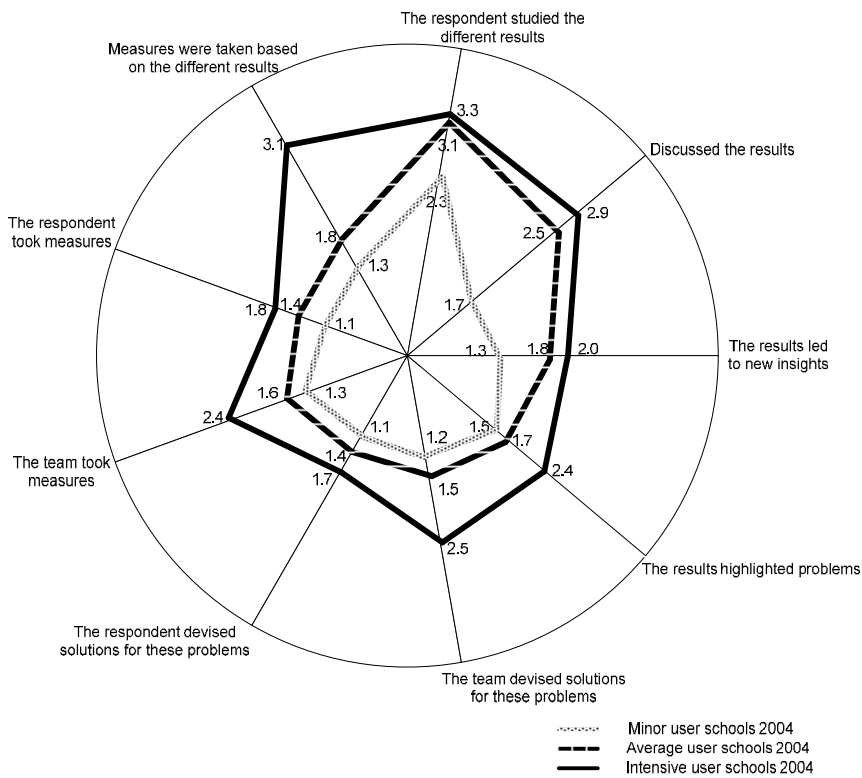


Figure 4.3 Mean scores of the LoSE, AvSE and HiSE schools on each of the ZEBO use variables in 2004 (variables measured on a four-point scale 1 = not at all, and 4 = to a great degree)

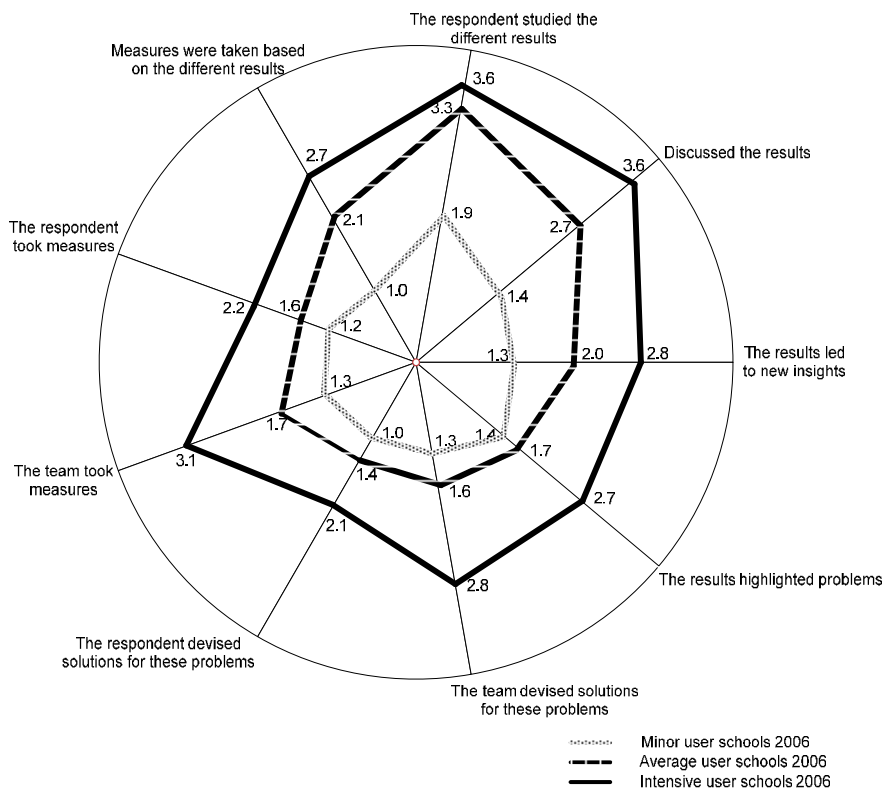


Figure 4.4 Mean scores of the LoSE, AvSE and HiSE schools on each of the ZEBO use variables in 2006 (variables measured on a four-point scale: 1 = not at all, and 4 = to a great degree)

4.3 ZEB0 Use: The Results from the Interviews in 2003 and 2005

Interviews were held to help interpret and explain the quantitative findings on ZEB0 use from the Evaluation of ZEB0 Questionnaire. Respondents from LoSE, AvSE, and HiSE schools were interviewed to study in depth how staff, within these different groups of schools, used the results from ZEB0. Using the interviews, it was possible to explore, for example, what was the focus of the discussion of ZEB0 output, in the LoSE, AvSE and HiSE schools, and how teachers and principals from the three groups of schools used this output to improve their functioning.

In 2003, thirty-one respondents were interviewed: 12 respondents from 4 LoSE schools, 15 respondents from 5 AvSE schools, and 4 respondents from 2 HiSE schools. In 2005 (using the 2004 questionnaire results on ZEB0 use as a starting point), 25 respondents were interviewed: 9 respondents from 3 LoSE schools, 7 respondents from 3 AvSE schools and 9 respondents from 3 HiSE schools (see Table 4.5).

Table 4.5 Number of respondents across the use groups interviewed in 2003 and 2005

Year	LoSE: respondents (schools)	AvSE: respondents (schools)	HiSE: respondents (schools)	Total
2003	12 (4)	15 (5)	4 (2)	31 (11)
2005	9 (3)	57 (3)	9 (3)	25 (9)

Only the major results from the interviews in 2003 and 2005 with the teachers and principals from LoSE, AvSE, and HiSE schools are presented here, more detailed results may be found in Appendix 4.1.

In most AvSE and HiSE schools ZEB0 output was studied and discussed, mostly in team meetings, but also in performance interviews, in the parents' council, in the participation council, and with the school advisory service. The discussions focussed on differences within the team, differences between the schools and the national mean, remarkable and extraordinary items, and items which required improvement. One teacher indicated (in 2005), for example:

We differed to a great extent in our ways of teaching. We have talked about this and we have tried to formulate a protocol, or guidelines for our way of working. We are trying to follow this protocol to get more similarity in our ways of teaching, and in our ways of thinking: why do you do the things you do, and what are guidelines of the team with respect to these things.

In 2005, the results from the 2003 ZEB0 use were compared with the results from the 2004 ZEB0 use to see whether or not differences existed. One principal stated, for example:

The differences between the team and the management have decreased. The team and the management are more on one line. That is great. We have worked towards this and it has paid off.

According to the respondents from the HiSE and AvSE schools, the use of ZEBO output, based on the second administration of ZEBO was easier, because the respondents knew what to expect and it did not feel like a one time only experience anymore.

Principals and teachers from both AvSE schools and HiSE schools indicated that some measures were taken based on the results from ZEBO in 2003 as well as in 2005, such as:

- § Implementing peer review;
- § Developing an action or priority list;
- § Participating in professional development courses;
- § Clarifying the school's objectives;
- § Using the results for making personal development plans;
- § Setting up a quality care project group.

Professional development received more attention in schools based on ZEBO use. A principal, for example, stated the following (in 2005):

We have connected them [the ZEBO output] with writing personal development plans. I formulated six questions: which results are striking in a positive or negative way, on which items do you want to work, what goals do you want to achieve, how do you want to achieve these goals, at what time do you want to have reached these goals, and how can another person observe the changes you have made during a classroom visit? The teachers used the ZEBO output to write their answers down.

It must be noted here, that in 2003, it was the principal, most of the time, who indicated that several measures had been taken. Principals, for example, developed action plans, used the information for the school plan, implemented classroom consultation, and compared the ZEBO output to a quality care instrument which had been used the previous year. Teachers confirmed in most cases that these actions had been taken. However, sometimes they were not aware of these measures. One principal, for example, used the ZEBO output to develop a school policy measure to reduce the workload of teachers, which had no effect at classroom level (so far). It is interesting that the teachers from this school were not aware of this change in school policy.

In 2005, teachers became more involved in the use of ZEBO. The following statement made by a teacher in 2005 illustrates:

We are really working with them [the ZEBO output] these days. With someone from the school advisory service we are studying the results. We are studying how these results came about and we are discussing whether we agree or disagree with the results. Do we want to change things, and how are we going to change these things? In project groups we all had to work out one of the

aspects. We had to start working on the items which scored below or above average. We had to ask ourselves the questions whether or not it was correct, how it came about, whether or not it was expected or surprising, whether or not we needed to change it, or whether it concurred with our vision? We are still working on this.

Teachers also made changes at the classroom level as the following citation indicates (2005):

I am trying to be clearer towards the children, by placing assignments on the black board, these kinds of things; and I am trying to take into account that certain children need less explanation.

Most striking from the interview results in 2003 is that three teachers from LoSE schools stated that they had never seen the results from ZEBO. This became clear from the following statement made by one of the teachers interviewed in 2003:

I am very negative about ZEBO, because we completed the questionnaires and never got the results back. Why did we fill in these questionnaires?

In general, in the LoSE schools, no measures were taken based on the output from ZEBO, with the exception of an action plan developed by one of the principals to improve weak aspects of the school's functioning, this plan, however, has not been implemented (yet).

In 2005, unlike in 2003, all respondents stated that they studied the ZEBO output. One principal and one of the other teachers said that the results did not concur with their expectations. In this school, problems existed between the school team and the principal. According to the teachers and the deputy head, ZEBO highlighted these problems. However, according to the principal the questions had been interpreted incorrectly and therefore no measures had been taken based on the ZEBO output. The principal stated:

We were enormously disappointed by the results, especially by the results from the teachers regarding the principal. I think the questions are difficult to interpret. I think we interpreted the questions wrongly.

However, the deputy head had a different opinion:

The results made us think, because there are clearly some communication problems. I had my suspicions all along. I had already picked up that there were problems in the communication with the principal. People sometimes even indicated that they preferred talking to me instead of talking to the principal.

4.4 A Comparison of the First, Second and Third Evaluations of ZEBO Use

Tables 4.6 and Table 4.7 show the comparison of the three evaluations of the use of ZEBO based on the results of the Evaluation of ZEBO Questionnaire. Table 4.6 shows the mean scores of the principals and teachers on the different ZEBO use variables in 2003, 2004, and 2006.

Table 4.6 *ZEBO use compared across the three evaluation years*

Variables	2003			2004			2006		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Conceptual use of ZEBO by principals	40	8.5	2.5	40	7.9	2.5	24	8.9	3.0
Conceptual use of ZEBO by teachers	120	6.5	2.8	147	6.5	2.4	101	7.0	2.5
New insights for principals	40	2.3	0.9	46	2.1	0.9	25	2.4	1.0
New insights for teachers	176	1.5	0.8	212	1.7	0.9	131	2.1	0.9
Instrumental use of ZEBO by principals	34	12.0	3.0	39	12.2	3.0	22	12.5	3.6
Instrumental use of ZEBO by teachers	85	9.5	3.5	104	9.7	3.7	75	11.7	3.3
The output was studied by the principals	41	3.6	1.0	48	3.7	0.8	25	3.4	1.0
The output was studied by the teachers	208	2.3	1.4	232	2.8	1.3	136	3.2	1.1
The output was discussed, according to principals	41	2.8	0.9	47	2.7	0.9	25	2.8	1.0
The output was discussed, according to teachers	196	2.3	1.0	227	2.4	1.0	138	2.7	1.0
The school took measures, based on the various ZEBO outputs, according to principals	37	1.9	1.2	43	2.1	1.4	23	2.0	1.6
The school took measures, based on the various ZEBO outputs, according to teachers	103	1.5	1.0	113	1.7	1.2	84	2.2	1.4

The results of the post hoc ANOVA test in Table 4.7 show which of the differences between the mean ZEBO scores for 2003, 2004, and 2006 is significant. Post hoc ANOVA tests encompass pair wise comparisons to compare all different combinations (Field, 2000). The use of ZEBO output by teachers and principals in 2003 was compared with the use in 2004, the use in 2004 to the use in 2006, and the use in 2003 was compared with the use of ZEBO output in 2006. Since a total of 36 comparisons were made, it was decided to choose a significance level of 0.01, to consider chance capitalisation. The significant increases in ZEBO use are printed in bold.

Table 4.7 Differences in ZEBO use over the years

Variables	Time (I)		Time (J)		Mean difference (J-I)	Std. error	Sig.
		N		N			
Conceptual use of ZEBO by principals	2003	40	2004	40	-0.65	0.60	0.52
	2004	40	2006	24	1.06	0.69	0.27
	2003	40	2006	24	0.42	0.69	0.82
Conceptual use of ZEBO by teachers	2003	120	2004	147	0.03	0.31	0.99
	2004	147	2006	101	0.52	0.33	0.25
	2003	120	2006	101	0.55	0.34	0.25
New insights for principals	2003	40	2004	46	-0.22	0.20	0.52
	2004	46	2006	25	0.29	0.23	0.41
	2003	40	2006	25	0.08	0.23	0.95
New insights for teachers	2003	176	2004	212	-0.20	0.10	0.09
	2004	212	2006	131	0.05	0.11	0.88
	2003	176	2006	131	-0.15	0.11	0.33
Instrumental use of ZEBO by principals	2003	34	2004	39	0.23	0.74	0.95
	2004	39	2006	22	0.28	0.84	0.94
	2003	34	2006	22	0.52	0.91	0.84
Instrumental use of ZEBO by teachers	2003	85	2004	104	0.25	0.52	0.88
	2004	104	2006	75	2.01	0.54	0.00
	2003	85	2006	75	2.26	0.56	0.00
The output was studied by the principals	2003	41	2004	48	0.13	0.19	0.79
	2004	48	2006	25	-0.33	0.24	0.38
	2003	41	2006	25	-0.20	0.27	0.73
The output was studied by the teachers	2003	208	2004	232	0.46	0.13	0.00
	2004	232	2006	136	0.39	0.13	0.00
	2003	208	2006	136	0.85	0.14	0.00
The output was discussed according to principals	2003	41	2004	47	-0.08	0.20	0.92
	2004	47	2006	25	0.08	0.23	0.94
	2003	41	2006	25	0.00	0.24	1.00
The output was discussed according to teachers	2003	196	2004	227	0.19	0.10	0.16
	2004	227	2006	138	0.27	0.11	0.03
	2003	196	2006	138	0.46	0.11	0.00
The school took measures based on the various ZEBO outputs according to principals	2003	37	2004	43	0.19	0.29	0.78
	2004	43	2006	23	-0.12	0.33	0.94
	2003	37	2006	23	0.08	0.34	0.97
The school took measures based on the various ZEBO outputs according to teachers	2003	103	2004	113	0.18	0.15	0.43
	2004	113	2006	84	0.51	0.19	0.02
	2003	103	2006	84	0.70	0.18	0.00

As Table 4.7 shows, instrumental use of ZEBO output by teachers increased significantly in 2006, as compared with 2003 and 2004. ZEBO output was studied and discussed significantly more, over the years according to teachers. Moreover, more measures were taken, based on the different kinds of ZEBO output, according to teachers over the years.

The use of ZEBO output by principals did not change significantly over the years. However, most principals were already using the ZEBO output at the start of this study. It also must be taken into account that over the years, several schools stopped participating in the project. Furthermore, twelve schools chose to postpone administering ZEBO to 2007. In 2003, for example, 41 principals completed the Evaluation of ZEBO Questionnaire, and in 2006 only 25 principals did this. All this means that the questionnaire results may present a slightly distorted view.

The differences are striking between principals and teachers. Almost half of the teachers indicated in the Evaluation of ZEBO Questionnaire from 2003, that they did not receive the ZEBO output, whereas most principals responded that they did receive and study these results. Principals in 2003 sometimes did not distribute the results among school staff. However, the 2004 and 2006 results show that the ZEBO output was distributed within the school, after the second and third evaluation of ZEBO use. In 2006, almost all teachers indicated, for example, that they studied the results.

Furthermore, it was usually the principal who indicated that measures were taken, based on the ZEBO output, to improve the quality of education. Most teachers confirmed that these measures were taken. However, some teachers were not aware of the fact that measures were taken, as a result of the use of ZEBO.

Although principals made more use of the ZEBO output, than teachers in 2003, 2004, and 2006, the results from 2004 and 2006 show that teachers are starting to catch up. A possible explanation for this increase in involvement may be that teachers start to realise that school self-evaluation is there to stay, and is not just another questionnaire they must fill out. Teachers become aware that self-evaluation information may be useful and helpful for their functioning and for their school.

The interview results confirm that teachers gradually became more involved in the use of ZEBO. In 2004 and 2006, all interviewed respondents indicated that they studied the results. Teachers also started to appreciate the ZEBO information, especially information from their pupils. Moreover, teachers and principals from the AvSE and HiSE schools compared the results of the first ZEBO use to the second ZEBO use. As for 2003, the interview results indicate that teachers started experiencing ZEBO as part of the regular school routine, as evidenced by the following statements: “it did not feel like a one time only experience anymore” and “the first ZEBO gave us a rough expression. The second time we have extracted specific items on which we want to improve”.

4.5 Conclusions

The results show that starting self-evaluation is not what causes difficulties. Schools did not experience difficulties in administering ZEBO. However, using performance feedback wisely for decisions in schools is much more than just gathering data and turning them into numbers. The data must be interpreted and translated into workable knowledge (Earl and Fullan, 2003). The steps after data collection prove to be the most difficult. After interpretation of the data, and after school staff acknowledge that problems exist, it takes time, money and effort to solve these problems. It is much easier to ignore the results and to continue in the old (and comfortable) way. Providing schools with information on their functioning is an insufficient stimulus for triggering improvement-oriented behaviour. Furthermore, the ZEBO system (and self-evaluation) is new to schools; it takes time to establish such systems within a school, and to become familiar with the system and the data it provides. For a considerable proportion of the schools it may still take several years before they have the knowledge, time and competencies needed to use self-evaluation results effectively.

From the 2003 results, it became clear that the conceptual and instrumental use of ZEBO were limited, especially use by teachers. A possible explanation of the limited conceptual ZEBO use may be that some teachers never received the ZEBO output, as became clear from the interviews. Almost half of teachers (44%) indicated, in the Evaluation of ZEBO Questionnaire, that they did not study the ZEBO output or did so to a minimal degree only. In most cases, principals controlled the ZEBO administration and the ZEBO output and therefore they did see the results. In one of the case studies on the use of the school performance feedback from the ABC+ model described by Teddlie et al (2002), something similar happened; only the principals received the feedback. Our findings are also in line with Gray (2002), who concludes that the use of pupil performance feedback is usually top-down. The principal controls access to the analyses and decides whether to share the information, or not. As Van Petegem and Vanhoof (2002b, 2004) state, one of the ways for principals to “use” school reports is to withhold these reports. If the evaluation results do not fit with the discourse of the principal, then principals may sometimes obscure and withhold the results, because the data would put a damper on the engagement of their teachers, and/or the data do not match their own experiences and information from other sources on their school’s functioning. The interview results also show that if a principal does not agree with the ZEBO information the information may sometimes be disregarded.

For instrumental ZEBO use, the results were similar to the results for conceptual use of ZEBO. Respondents from 24% of the schools (n=50), who worked with ZEBO, used the results instrumentally in 2003. These respondents indicated that the results were discussed. This is an important finding since dialogue and reflection are the first steps on the way to developing actions to improve institutional performance, which is the main goal of ZEBO use. Moreover, respondents from these schools mentioned that actions were taken to improve the quality of education. Measures which were taken included improving communication, stimulating independent learning, more differentiation, implementation of classroom consultation, and more frequent evaluation and testing of pupils.

The conceptual use of ZEBO was still limited in 2004. Fewer principals reported that the results from ZEBO led to new insights in 2004. Schools used ZEBO for the second time, and principals probably knew what to expect. Teachers reported a higher number of new insights. This may be due to the fact that in 2003 almost half of the teachers never saw the ZEBO output, whereas in 2004, 61% of teachers (n= 236) who completed the Evaluation of ZEBO Questionnaire, indicated that they had studied the ZEBO output.

Instrumental use was also still limited 2004. Thirteen schools (26%) used ZEBO output instrumentally. The ZEBO output was discussed more. Teachers reported that ZEBO output provided a common starting point for discussion. Extensive discussions may help school staff to arrive at a deeper understanding of the ZEBO output and of the implications for their work. Discussion may give meaning to the data and shared values on performance feedback and school improvement may originate from discussions

(Geijsel & Krüger, 2005). The results from the second evaluation of the use of ZEBOS further show that the teachers and principals from the schools which used the ZEBOS output instrumentally indicated that the school team took measures based on to improve the quality of education.

In 2006, (most) schools used ZEBOS for the third time. School staff used the output more than in 2003, and 2004. Although it must be noted here that only 43 schools administered ZEBOS, compared with 64 in 2003 and 58 schools in 2004.

At the start of this study not even half of the teachers had received ZEBOS output. Teachers were simply not involved in the school self-evaluation process. They only completed the ZEBOS questionnaire. In 2006, the majority of teachers (71%) had seen the ZEBOS output and as a result they were able to make changes in their classrooms. Teachers indicated, in the Evaluation of ZEBOS Questionnaire and interviews, that they accommodated to differences more, made their education more adaptive, and brought more clarity in their lessons. Principals used the information for making changes at school level, for example, for writing the school plan, for school policy development, and for conducting performance interviews.

Overall, ZEBOS use has increased slightly but significantly over the years. The fact that evaluation use is a phenomenon which gradually develops over time was also noted by Saunders (2000). She found, in a study into the use of value-added data, that schools which were participating in the project for at least three years, were more likely to be making active use of the performance data. School staff in this ZEBOS evaluation study made more use of ZEBOS output over the years, but most schools still have a long way to go. Although in 2006, 24% of the respondents used the results instrumentally, 66% of the respondents did not. This is in line with other studies (Weiss, 1998a, 1998b; Coe & Visscher, 2002b; Van Petegem & Vanhoof, 2002a, 2002b; Van Petegem & Vanhoof, 2004) which show that in many schools, performance feedback is under-utilised, valuable information is not used, or only to a small degree.

Engaging in a school self-evaluation process remains difficult. Most teachers are busy with their day-to-day classroom teaching tasks, pupil evaluation tasks, and school administration tasks. School self-evaluation is another task added to their work. As a result they may have little time to contemplate on the school self-evaluation results, and how to use these effectively. They may also lack the necessary skills and competencies to use the output effectively (Ghere, King, Stevahn & Minnema, 2006; Clift et al., 1987).

The under-utilisation of the ZEBOS evaluation results, in most schools, will have its implications for the degree to which improvement oriented actions actually happen and ultimately also, for outcomes, in the sense of improved pupil performance. This issue will be taken up in the next chapter.

Chapter 5

The Effects of ZEBO use

5.1 Introduction

The first objective of this chapter is to clarify whether ZEBO use had an effect on pupil achievement. These effects are likely to be mediated through school organisational and teaching processes, such as educational leadership, better communication and more effective teaching strategies. The effect of ZEBO use on these school and classroom process indicators was also studied. Various studies into the effects of the use of School Performance Feedback Systems (Teddlie et al., 2002; Hendriks et al., 2002; Gray, 2002; Rowe et al., 2002; Tymms & Albone, 2002) indicate some positive effects on specific prerequisites for improved school performance, such as, an effect on didactic behaviour, an increase in professional development activities and an improved functioning of the principal (Davies & Rudd, 2002; Webb et al., 1998). Finally, it cannot be ruled out that the use of a self-evaluation instrument has unintended consequences, such as increased workload or stress. The possible occurrences of these negative, unintended effects were therefore, also studied.

The first question to be answered is whether ZEBO use had an effect on pupil achievement. To study the effect of ZEBO use on pupil achievement, the test scores for spelling (SVS) and mathematics (Maths) were collected, before schools administered ZEBO for the first time, in June 2002. Test scores were again obtained in the year 2003 (group 4 and 5), 2004 (group 5 and 6), 2005 (group 6 and 7), and 2006 (group 7 and 8). Test scores were obtained twice a year. Two cohorts of pupils were followed:

- § Cohort 1: pupils from grade 3 (age 6) to grade 7 (age 11);
- § Cohort 2: pupils from grade 4 (age 7) to grade 8 (age 12).

The possible relationship between the conceptual and instrumental use of ZEBO with spelling and mathematics achievement was analysed per cohort. Multilevel analyses, with repeated measures, were conducted, since data from five years had been collected (more information on the data analysis design may be found in Chapter 3).

The second question explored, in this chapter, is what, according to principals and teachers, were the (intended and unintended) ZEBO use effects. To answer this question, the Evaluation of ZEBO Questionnaire was administered three times; each time after schools had administered ZEBO. In the questionnaire, respondents were asked to what degree they felt that ZEBO use had an effect on the process indicators shown in Figure 5.1. Figure 5.1 displays the variables associated with the effects of ZEBO use.

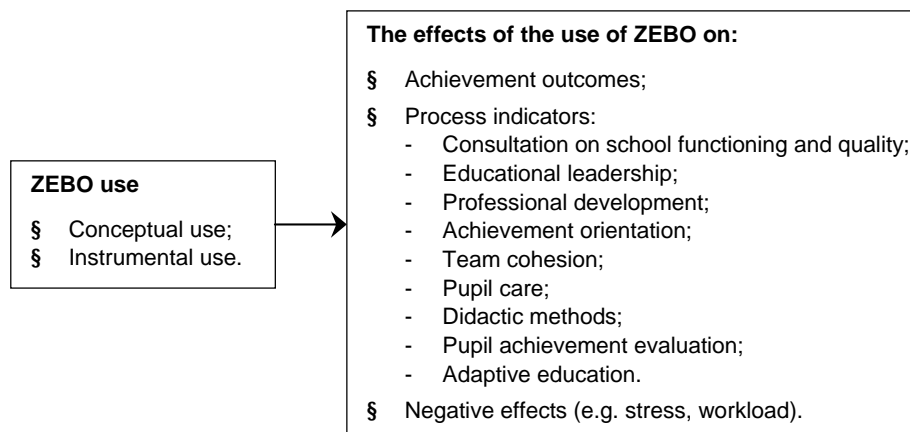


Figure 5.1 The variables used to explore the effects of ZEBU use

The effect of ZEBU use on pupil achievement is presented in section 5.2, followed by a conclusion in section 5.3. Thereafter, in section 5.4 the perceived effects on process indicators (see fig. 5.1) are presented. In section 5.5, the results of the evaluation of the perceived effects of ZEBU use in 2003, 2004, and 2006 are compared with each other. The chapter ends with conclusions regarding the perceived effects of ZEBU use (section 5.6).

5.2 Effect of ZEBU Use on Pupil Achievement: Multilevel Analyses with Repeated Measures

The data in this study were longitudinal, repeated measures, which may be viewed as multilevel data, with repeated measures nested within individuals (Hox, 2002). In the case of repeated measures, the same individuals are measured on more than one occasion.

An advantage of multilevel modelling with repeated measures is that it allows a repeated measures variable, such as time, to be treated as random, and nested within upper-level units, such as pupils and schools (Snijders & Bosker, 1999; Rasbash et al., 2000; Hox, 2002).

Another advantage of conducting multilevel analyses with repeated measures is that these analyses do not require the same number of measurement occasions per individual subject. Although some pupils had left the school, repeated a class, or missed a test, it is still possible to incorporate all available data into the analyses (Snijders & Bosker, 1999; Rasbash et al., 2000; Hox, 2002). In MANOVA, usually cases with missing measurements are removed from the analysis. Multilevel regression models do not assume an equal number of observations, so cases with missing measurements may remain in the analysis, which results in a larger sample, and larger samples increase the precision of the estimates and the power of statistical tests (Hox, 2002).

A further benefit of multilevel analyses, with repeated measures, is that the repeatedly measured response (in this study, pupil achievement) may be modelled over time, as a continuous growth curve, instead of as a series of interval measurements. By modelling regression coefficients at the occasion level, growth curves which are different for each individual pupil may also be estimated (Snijders & Bosker, 1999; Rasbash et al., 2000; Hox, 2002).

Two cohorts of pupils were followed as described above. For both cohorts multilevel analyses with repeated measures were conducted to measure the effects of ZEBO use on pupil achievement growth. Pupil achievement is the response, and there are three levels of data: schools (level 3), pupils (level 2) and measurement occasion (level 1). In addition, there are two explanatory variables. The first is the conceptual use of ZEBO (in 2003, 2004, and 2006). The other explanatory variable is the instrumental ZEBO use (in 2003, 2004, and 2006). Various pupil characteristics were taken into account as covariates, because significant differences between pupils on these characteristics exist¹.

For the multilevel analysis with repeated measures in 2006, the results of the first cohort of pupils (pupils who started in grade 3 and who were in grade 7 in 2006) on the standardized Cito spelling and mathematics tests were used. Pupils' spelling attainment and mathematics attainment were tested on up to nine occasions: twice a year from 2002 to 2006. Only the data from the pupils from schools which had administered ZEBO at least twice, were used for the multilevel analysis: 3.839 observations from 474 pupils.

Almost all pupils from the second cohort (pupils who started in grade 4 and who were in grade 8 at the time of the third ZEBO administration) took tests other than the standardized Cito spelling and mathematics tests in 2006, and therefore no data could be collected for the grade 8 pupils (cohort 2) in 2006. Therefore, to study the effect of ZEBO use on pupil achievement growth for the second cohort, multilevel analyses with repeated measures in 2005 (instead of in 2006) were conducted. Pupils' spelling and mathematics attainment levels were tested on up to seven occasions: twice a year from 2002 to 2005. For the second cohort, 7.308 observations from 1.044 pupils were obtained.

Table 5.1 shows the analyses conducted in order to study the effect of ZEBO use on pupil achievement growth over the years, including the dependent and independent variables.

¹ At the beginning of the school year 2003/2004 teachers were asked to fill out a pupil form to gather information about the following pupil background characteristics which may influence pupil achievement: gender, whether or not the pupil resided in a so-called combination class (in which pupils from more than one grade level are taught in one class), socio economic status (SES) (measured with pupil "weight"), language at home (Dutch, dialect, Turkish or other), perceived intelligence (low, average, or high IQ), class size, and age.

Table 5.1 The multilevel analyses for cohorts 1 and 2.

Year	Multilevel analyses with repeated measures for		Independent variable
		Dependent variable	
2006	Cohort 1	Growth in <i>spelling</i> achievement. Measured on 9 occasions: grade 3 (1 test), grade 4 (2 tests), grade 5 (2 tests), grade 6 (2 tests), grade 7 (2 tests)	Conceptual ZEBO use in 2003, 2004, and 2006
	Cohort 1	Growth in <i>spelling</i> achievement. Measured on 9 occasions: grade 3 (1 test), grade 4 (2 tests), grade 5 (2 tests), grade 6 (2 tests), grade 7 (2 tests)	Instrumental ZEBO use in 2003, 2004, and 2006
	Cohort 1	Growth in <i>mathematics</i> achievement. Measured on 9 occasions: grade 3 (1 test), grade 4 (2 tests), grade 5 (2 tests), grade 6 (2 tests), grade 7 (2 tests)	Conceptual ZEBO use in 2003, 2004, and 2006
	Cohort 1	Growth in <i>mathematics</i> achievement. Measured on 9 occasions: grade 3 (1 test), grade 4 (2 tests), grade 5 (2 tests), grade 6 (2 tests), grade 7 (2 tests)	Instrumental ZEBO use in 2003, 2004, and 2006
2005	Cohort 2	Growth in <i>spelling</i> achievement. Measured on 7 occasions: grade 4 (1 test), grade 5 (2 tests), grade 6 (2 tests), and grade 7 (2 tests)	Conceptual ZEBO use in 2003 and 2004
	Cohort 2	Growth in <i>spelling</i> achievement. Measured on 7 occasions: grade 4 (1 test), grade 5 (2 tests), grade 6 (2 tests), and grade 7 (2 tests)	Instrumental ZEBO use in 2003 and 2004
	Cohort 2	Growth in <i>mathematics</i> achievement. Measured on 7 occasions: grade 4 (1 test), grade 5 (2 tests), grade 6 (2 tests), and grade 7 (2 tests)	Conceptual ZEBO use in 2003 and 2004
	Cohort 2	Growth in <i>mathematics</i> achievement. Measured on 7 occasions: grade 4 (1 test), grade 5 (2 tests), grade 6 (2 tests), and grade 7 (2 tests)	Instrumental ZEBO use in 2003 and 2004

The influence of conceptual and instrumental ZEBO use on spelling and mathematics achievement was first analysed per cohort, three times, each year after schools had used ZEBO. These analyses were all associated with the effect of ZEBO use, on pupil achievement, in one year (2003, 2004 or 2006). The results of these analyses may be found in Appendices 5.1 to 5.10. The results of these analyses indicate that in one year, ZEBO use did not influence pupil achievement.

Schools were supposed to administer ZEBO three times². Therefore the ultimate question is whether the aggregate of ZEBO use, over the years, is related to pupil progress, over the years. To answer this question, multilevel analyses with repeated measures were conducted. The results are presented in the following section.

² However, not all schools administered ZEBO three times. Thirty-six schools administered ZEBO three times, twenty-two schools two times, and 10 schools administered ZEBO only one time.

5.2.1 The Effect of ZEBO Use on Pupil Achievement Growth for Cohort 1

This section seeks to answer the following research questions for cohort 1 (pupils who started in grade 3):

1. Do pupils from schools which score more highly on *conceptual* (see Figure 5.1) ZEBO use (as measured by the Evaluation of ZEBO Questionnaire in 2003, 2004 and 2006) make more progress over time on spelling than pupils from schools with lower scores on *conceptual* ZEBO use?
2. Do pupils from schools which score more highly on *conceptual* ZEBO use (as measured by the Evaluation of ZEBO Questionnaire in 2003, 2004 and 2006) make more progress over time on mathematics than pupils from schools with lower scores on *conceptual* ZEBO use?
3. Do pupils from schools which score more highly on *instrumental* (see Figure 5.1) ZEBO use (as measured by the Evaluation of ZEBO Questionnaire in 2003, 2004 and 2006) make more progress over time on spelling than pupils from schools with lower scores on *instrumental* ZEBO use?
4. Do pupils from schools which score more highly on *instrumental* ZEBO use (as measured by the Evaluation of ZEBO Questionnaire in 2003, 2004 and 2006) make more progress over time on mathematics than pupils from schools with lower scores on *instrumental* ZEBO use?

The answer to these questions lies in the value and significance of the parameter which indicates the interaction effect of (1) the time variable with (2) the ZEBO use variable. The multilevel analyses for cohort 1 were first conducted with *spelling* achievement as the dependent variable. Next, analyses were conducted with *mathematics* achievement as the dependent variable (see Table 5.1). In both analyses, the aggregate conceptual ZEBO use was entered as an independent variable. The aggregate conceptual use variable was constructed by taking the conceptual use of ZEBO in 2003, 2004, and 2006 together. This aggregate variable, calculated for each school, was entered into the multilevel analyses as a continuous variable. Analyses with instrumental ZEBO use, as the independent variable, were also conducted with spelling achievement as the dependent variable, and mathematics achievement as the dependent variable. By adding the sum of instrumental use of ZEBO per school in 2003, 2004, and 2006, an aggregate instrumental use variable was calculated.

The results of the multilevel analyses conducted for cohort 1 are presented in Table 5.2, which shows that the growth in pupil spelling achievement was influenced by IQ and gender. Pupils with a higher IQ, and girls, made more progress over time on spelling than boys, or pupils with a lower IQ. These pupil characteristics were therefore taken into account in the analyses. The model with (1) the added ZEBO conceptual use variable, and (2) the interaction between conceptual ZEBO use and time, did not fit the data significantly better than the basic model without the ZEBO use variables, nor was the improvement fit for the conceptual use model ($\chi^2 = 0.02$, $df=2$, $p=0.99$) significant.

This means, in general, that pupils from schools which scored more highly on conceptual use of ZEBO (e.g. made more use of the ZEBO output) did not make more progress over time on spelling than pupils from schools with lower scores on conceptual ZEBO use. Nor is there an effect of instrumental ZEBO use on spelling growth. The improvement fit for the instrumental use Model ($\chi^2 = 0.15$, $df=2$, $p=0.93$) was not significant.

Table 5.2 *Multilevel repeated measures analyses of the effect of conceptual ZEBO use and instrumental ZEBO use over the years on growth in spelling achievement for cohort 1*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	105.69 (2.78)	106.79 (2.83)
Time	4.24 (0.40)	4.01 (0.42)
Av. vs. low IQ	3.84 (0.54)	3.85 (0.54)
High vs. low IQ	9.69 (0.62)	9.68 (0.62)
Girls vs. Boys	2.39 (0.34)	2.40 (0.34)
Conceptual ZEBO use	0.01 (0.11)	
Time*conceptual ZEBO use	-0.01 (0.02)	
Instrumental ZEBO use		-0.02 (0.07)
Time*instrumental ZEBO use		0.00 (0.01)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	6.32 (1.62)	6.29 (1.62)
σ^2_{v1}	-0.66 (0.20)	-0.66 (0.20)
Level 2 (pupil)		
σ^2_{u0}	16.23 (1.29)	16.24 (1.29)
σ^2_{u1}	0.42 (0.14)	0.42 (0.14)
Level 3 (occasion)		
σ^2_e	21.36 (0.43)	21.36 (0.43)

Note: Standard errors between brackets. * indicates interaction effect.

Secondly, it was hypothesized that pupils from schools which score more highly on conceptual or instrumental use of ZEBO, would make more progress over time on mathematics, than pupils from schools with lower scores on conceptual or instrumental ZEBO use. The results indicate that a more intensive use of ZEBO did not result in pupils making more progress over time on mathematics either (see Table 5.3).

Table 5.3 *Multilevel repeated measures analyses of the effect of conceptual ZEBO use and instrumental ZEBO use over the years on growth in mathematics achievement for cohort 1*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	58.20 (4.46)	55.58 (4.69)
Time	5.98 (0.56)	6.63 (0.58)
Av. vs. low IQ	6.40 (0.81)	6.41 (0.81)
High vs. low IQ	14.67 (0.94)	14.67 (0.94)
Girls vs. Boys	-2.55 (0.49)	-2.55 (0.49)
Number of pupils	0.29 (0.11)	0.28 (0.11)
SES (1.25 vs. 1.0)	-2.19 (0.70)	-2.20 (0.70)
SES (1.9 vs. 1.0)	-5.33 (1.01)	-5.37 (1.01)
Conceptual use	-0.16 (0.15)	
Time*conceptual use	0.02 (0.56)	
Instrumental use		-0.03 (0.10)
Time*instrumental use		0.00 (0.02)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	10.56 (3.00)	10.90 (3.07)
σ^2_{v1}	-0.73 (0.36)	-0.78 (0.37)
Level 2 (pupil)		
σ^2_{u0}	49.39 (3.31)	49.39 (3.31)
σ^2_{u1}	-1.39 (0.33)	-1.39 (0.33)
Level 3 (occasion)		
σ^2_e	37.82 (0.76)	37.82 (0.76)

Note: Standard errors between brackets, * means interaction effect.

5.2.2 The Effect of ZEBO Use on Pupil Achievement Growth for Cohort 2

Almost all pupils from the second cohort (pupils who started in grade 4 and who were in grade 8 during the third ZEBO administration) took tests other than the standardized Cito spelling and mathematics tests, and for this reason no data could be collected for the grade 8 pupils (cohort 2) in 2006. Therefore, to study the effect of ZEBO use on pupil achievement growth for the second cohort, the analyses were conducted with pupil achievement scores from grade 4 (2002) to grade 7 (2005) as the dependent variable, and ZEBO use in 2003 and 2004 as independent variable.

This section seeks to answer the following research questions for cohort 2:

1. Do pupils in schools which score more highly on *conceptual* use of ZEBO (as measured by the Evaluation of ZEBO Questionnaire in 2003 and 2004) make more progress over time on spelling than pupils in schools with lower scores on *conceptual* ZEBO use?
2. Do pupils in schools which score more highly on *conceptual* use of ZEBO (as measured by the Evaluation of ZEBO Questionnaire in 2003 and 2004) make more progress over time on mathematics than pupils in schools with lower scores on *conceptual* ZEBO use?
3. Do pupils in schools which score more highly on *instrumental* use of ZEBO (as measured by the Evaluation of ZEBO Questionnaire in 2003 and 2004) make more progress over time on spelling than pupils in schools with lower scores on *instrumental* ZEBO use?
4. Do pupils in schools which score more highly on *instrumental* use of ZEBO (as measured by the Evaluation of ZEBO Questionnaire in 2003 and 2004) make more progress over time on mathematics than pupils in schools with lower scores on *instrumental* ZEBO use?

The multilevel analyses for cohort 2 were also conducted with spelling and mathematics achievement as the dependent variables. The aggregate conceptual use variable was calculated by taking the conceptual use of ZEBO in 2003 and 2004 together. By adding the instrumental use of ZEBO in 2003 and 2004, an aggregate instrumental ZEBO use variable was calculated.

The results in Table 5.4 show that the growth of pupil spelling achievement in the second cohort was also influenced by IQ and gender. Pupils with a higher IQ and girls made more progress over time on spelling than boys and pupils with a lower IQ. These pupil characteristics were therefore taken into account in the analyses. Pupils from schools with more intensified conceptual use of ZEBO did not make more progress on spelling than pupils from schools with less intensified conceptual ZEBO use. The model with (1) the ZEBO conceptual use variable and (2) with the interaction effect between conceptual ZEBO use and time, did not fit the data significantly better than the basic model without the ZEBO use variables, nor was the improvement fit for the conceptual use Model ($\chi^2 = 0.73$, $df=2$, $p=0.69$) significant.

Instrumental use of ZEBO was not associated with pupils making more progress over time on spelling or mathematics either. The improvement fit for the instrumental use model ($\chi^2 = 1.89$, $df=2$, $p=0.39$) was not significant.

Table 5.4 *Multilevel repeated measures analyses of the effect of conceptual ZEBO use and instrumental ZEBO use over the years on growth in spelling achievement for cohort 2*

	Conceptual use Model	Instrumental use Model
<i>Fixed</i>		
Intercept	115.52 (1.34)	116.39 (1.44)
Time	3.78 (0.22)	3.75 (0.24)
Av. vs. low IQ	3.86 (0.45)	3.88 (0.45)
High vs. low IQ	9.33 (0.51)	9.34 (0.51)
Girls vs. Boys	2.19 (0.30)	2.19 (0.30)
Conceptual ZEBO use	0.03 (0.07)	
Time*conceptual ZEBO use	-0.01 (0.01)	
Instrumental ZEBO use		-0.02 (0.06)
Time*instrumental ZEBO use		-0.01 (0.01)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	7.45 (1.67)	7.44 (1.66)
σ^2_{v1}	-1.17 (0.27)	-1.18 (0.27)
Level 2 (pupil)		
σ^2_{u0}	19.50 (1.04)	19.50 (1.04)
σ^2_{u1}	0.00 (0.00)	0.00 (0.00)
Level 3 (occasion)		
σ^2_e	22.65 (0.42)	22.65 (0.42)

Note: Standard errors between brackets. * indicates interaction effect.

It was also hypothesized that pupils from schools which score more highly on conceptual or instrumental use of ZEBO, would make more progress over time on mathematics achievement, than pupils from schools with lower scores on conceptual or instrumental ZEBO use. Analyses for the mathematics tests are presented in Table 5.5. From the results, it may be concluded that ZEBO use did not result in pupils from cohort 2 making more progress over time on mathematics either³.

³ All analyses for cohort 1 and cohort 2 were also conducted using a dummy ZEBO use variable (1=low use, 2=average use, 3=high use) to study whether the pupils from the Low Self-Evaluation (LoSE) schools made less progress on spelling or mathematics achievement over time than the pupils from the High Self-evaluation (HiSE) schools. These analyses also revealed no significant relationships between ZEBO use and pupil achievement. Furthermore, analyses were conducted using a variable representing the increase or decrease in the conceptual use of ZEBO, or the instrumental use of ZEBO. These analyses also revealed no significant relationships between ZEBO use and pupil achievement.

Table 5.5 *Multilevel repeated measures analyses of the effect of conceptual ZEBO use and instrumental ZEBO use over the years on growth in mathematics achievement for cohort 2*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	70.35 (2.21)	72.01 (2.32)
Time	5.91 (0.32)	5.62 (0.35)
Av. vs. low IQ	8.11 (0.62)	8.13 (0.62)
High vs. low IQ	16.38 (0.72)	16.39 (0.72)
Girls vs. Boys	-3.07 (0.41)	-3.08 (0.41)
Number of pupils	0.19 (0.05)	0.18 (0.05)
SES (1.25 vs. 1.0)	-2.48 (0.57)	-2.48 (0.57)
SES (1.9 vs. 1.0)	-4.62 (0.88)	-4.68 (0.88)
Conceptual use	0.01 (0.09)	
Time*conceptual use	0.01 (0.02)	
Instrumental use		-0.07 (0.08)
Time*instrumental use		0.02 (0.02)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	12.14 (2.76)	11.89 (2.72)
σ^2_{v1}	-2.18 (0.51)	-2.12 (0.50)
Level 2 (pupil)		
σ^2_{u0}	36.57 (1.92)	36.58 (1.92)
σ^2_{u1}	0.00 (0.00)	0.00 (0.00)
Level 3 (occasion)		
σ^2_e	38.39 (0.71)	38.39 (0.71)

Note: Standard errors between brackets, * means interaction effect.

5.3 Effects of ZEBO Use on Pupil Achievement: Conclusion

The results of the various multilevel analyses conducted in order to explore the relationships between conceptual use and instrumental use of ZEBO, and pupil achievement revealed no significant relationships. Although ZEBO use increased over the years, it remained still very limited (the respondents from only 13 of 31 schools used the ZEBO output instrumentally in 2006, as mentioned in Chapter 4). This limited use may explain why no effects of ZEBO use on pupil achievement were found.

Along with the effect of ZEBO use on pupil achievement, it was also interesting to investigate to what degree improvement on important school and classroom process indicators may be observed as potential effects of ZEBO use. ZEBO provides schools with information on process variables which have been shown to be associated with relatively high pupil achievement levels. The next section explores, whether according to principals and teachers, schools improved on these process indicators, as a result of ZEBO use.

5.4 Perceived Effects of Using ZEBO: Results of the Evaluation of ZEBO Questionnaire

This section seeks to answer the question, what are the perceived (intended and unintended) effects of ZEBO use on the process indicators, shown in Figure 5.2. To answer this question, schools received the Evaluation of ZEBO Questionnaire after each

time they administered ZEBU (in 2003, 2004 and 2006). Teachers and principals were asked to rate to what degree ZEBU use had an effect on the process indicators in Figure 5.2. The variables were measured on a four-point scale: 1= to a minimal degree/not, and 4= to a great degree. In addition, respondents were asked to provide concrete examples of perceived effects on the process indicators.

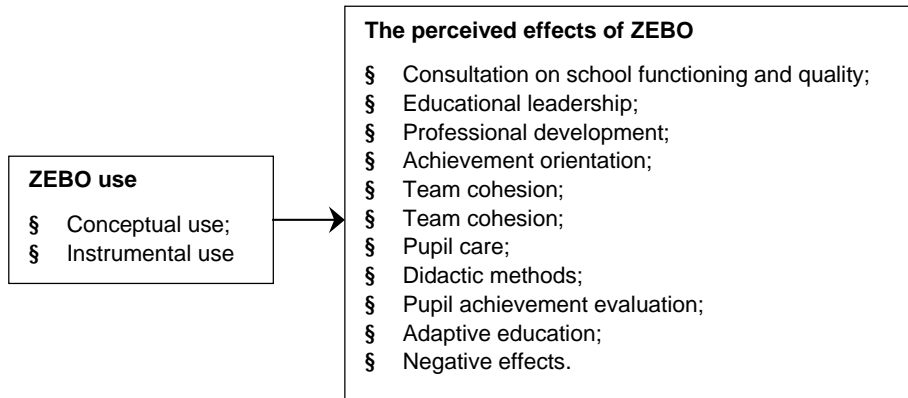


Figure 5.2 The perceived effects of ZEBU use

The results of the first evaluation of the effects of ZEBU on process indicators (in 2003) are presented below. For the second (2004) and third (2006) evaluation a summary of the results is presented.

5.4.1 Results of the Evaluation of ZEBU Questionnaires in 2003

Table 5.6 shows the mean scores of principals and teachers on the process indicators studied in 2003, 2004, and 2006. Table 5.6 shows that, consistent with the limited use of the ZEBU output, the perceived effects on process indicators were also limited. The mean score is equal or higher than 2, which means that the respondent said that the effects in question were perceived as being present at least “to a small degree“ on only two variables: “consultation on school functioning and quality according to principals” and “professional development according to principals”. The exact ratings on the process variables from the respondents on the four point-scales may be found in Appendix 5.11.

Table 5.6 *The perceived effects of ZEBO use in 2003, 2004, and 2006 on process indicators.*

Variables	2003			2004			2006		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Consultation on school functioning and quality according to principals	40	2.1	0.9	46	2.2	0.8	25	2.5	0.9
Consultation on school functioning and quality according to teachers	186	1.6	0.9	220	1.8	0.9	130	2.0	0.8
Educational leadership according to principals	37	1.8	0.7	41	1.7	0.8	21	1.9	0.8
Educational leadership according to teachers	146	1.3	0.7	158	1.5	0.7	96	1.5	0.7
Professional development according to principals	39	2.0	0.8	46	1.9	0.8	25	2.0	0.9
Professional development according to teachers	169	1.3	0.6	194	1.4	0.7	118	1.6	0.8
Achievement orientation according to principals	39	1.7	0.9	45	1.8	0.8	23	2.0	0.9
Achievement orientation according to teachers	172	1.5	0.8	196	1.4	0.7	117	1.7	0.8
Team cohesion according to principals	38	1.6	0.7	38	1.6	0.8	23	1.6	0.6
Team cohesion according to teachers	179	1.3	0.7	195	1.3	0.7	116	1.5	0.8
Pupil care according to principals	40	1.7	0.9	44	1.8	1.0	23	1.8	0.8
Pupil care according to teachers	168	1.5	0.9	194	1.6	0.9	118	1.6	0.9
Didactic methods according to principals	30	1.3	0.7	45	1.5	0.7	19	1.8	0.6
Didactic methods according to teachers	174	1.2	0.5	202	1.3	0.5	125	1.5	0.7
Pupil achievement evaluation according to principals	36	1.4	0.8	43	1.7	0.9	23	1.6	0.7
Pupil achievement evaluation according to teachers	194	1.3	0.7	207	1.3	0.7	131	1.5	0.7
Adaptive education according to principals	39	1.4	0.7	41	1.5	0.7	24	1.6	0.5
Adaptive education according to teachers	182	1.4	0.8	197	1.4	0.7	124	1.6	0.8

When studying the exact ratings, it became clear that a relatively higher number of improvements in 2003 were reported for the variable “consultation on school functioning and quality”. Sixteen percent of the teachers (n=220) and thirty-four percent of the principals (n=41) indicated that, in their view, ZEBO use led to more consultation on school functioning and quality “to a moderate degree”. As far as examples are concerned, principals and teachers both mentioned the agreement to discuss the quality of education, at least twice a year, in formal team meetings.

Twelve percent of teachers (n=220) in 2003 also indicated that ZEBO use led to an improvement of the pupil care system. Examples of improved pupil care included faster detection of special need pupils, developing a plan of action to improve the special needs support structure, and an increase in attention to the socio-emotional development of pupils. Furthermore, some principals (n=41) and teachers (n=25) reported that ZEBO had an effect on the following (to a moderate degree):

- § Professional development, 24% of the principals and 6% of the teachers mentioned, for example, that they took professional development courses;
- § The achievement orientation of the school, 20% of the principals and 11% of the teachers mentioned as an example here, the analysis of the teaching materials and methods used, and the replacement of some of the materials and methods;
- § Adaptive education, 15% of the principals and 11% of the teachers indicated, for example, that because of the measures taken to stimulate independent learning, teachers could now spend more time on special need pupils.

Principals and teachers, in general, did not report strong negative effects. Only one principal and eight teachers indicated that ZEBO led to negative effects. The negative effects which were mentioned included an increase in workload, and in a few cases, stress, this tended to be associated with the fact that the respondent did not agree with the ZEBO output.

5.4.2 Results of the Evaluation of ZEBO Questionnaires in 2004 and 2006

ZEBO use in 2004 and 2006 had similar effects as in 2003. The exact ratings on the process variables may be found in Appendix 5.11. Most improvements were again found for the variable “consultation on school functioning and quality”. Examples of this effect included:

- § Discussing and improving the mission statement of the school;
- § More consultation to develop improvement plans for the school;
- § An increase in the number of grade level meetings.

Professional development was also perceived as having improved in 2004 and 2006. In 2006, professional development improved to a moderate degree, according to 17% of the teachers (n=141), whereas in 2003 only 6% of the teachers (n=220) mentioned an improvement on this variable. The ZEBO output was, for example, used for writing professional development plans.

Teachers (17%, n=141) and principals (20%, n=25) responded in 2006 that the achievement orientation of school staff increased as a result of ZEBO, for example, in one case new instructional methods were acquired. Another example was that schools started to put on record, the achievement results of each grade, meticulously.

In 2006, more teachers, as compared with 2003 and 2004, indicated that ZEBO use had an effect on their didactic behaviour (11%, n=141) (e.g. pupils were more involved in the lessons and were more challenged by the teachers).

Using ZEBO was, for school staff from several schools, a positive experience from which they benefited, as the following statements demonstrate: “because of ZEBO use issues became discussable”, “before we started using ZEBO, we only had the inspection report as evaluation information; now we have something to compare it with”, and “ZEBO helps in developing and improving school policy”.

Principals and teachers did not report strong negative effects in 2003 nor did they do so in 2004 and 2006.

Based on the results of the Evaluation of ZEBO Questionnaire in 2003, 2004, and 2006, the schools which worked with ZEBO were divided into three groups based on the total score for the ZEBO use variables (for further details see Chapter 3):

- § One group of schools which did not use the ZEBO output (Low Self Evaluation, “LoSE”);
- § One group of schools which made average use of the ZEBO output (Average Self Evaluation, “AvSE”);
- § One group of schools which used the ZEBO output intensively (High Self Evaluation, “HiSE”).

In 2003, the respondents from the LoSE and AvSE schools reported almost no effects at all. The respondents from the HiSE schools reported some minor effects of ZEBO use on the process indicators. For the variables “consultation on school functioning and quality”, “the achievement orientation”, and “pupil care”, these schools scored between 2.0 and 2.3 on a four-point scale.

In 2004, the LoSE schools and AvSE schools again indicated that ZEBO use had almost no effects at all. HiSE schools reported some minor effects on “consultation on school functioning and quality” and “adaptive education” (respectively 2.3 and 1.9).

In 2006 ZEBO use was found to have more perceived effects compared with 2003 and 2004. The mean scores from the LoSE (dotted line), the AvSE (line with stripes) and the HiSE (bold line) schools in 2006 on the different effect variables are shown in Figure 5.3. As in 2003 and 2004, the LoSE schools perceived (almost) no effects of ZEBO on the process indicators in question. In the AvSE schools ZEBO use also had small perceived effects, although some respondents mentioned an increase in consultation on the quality of school functioning and quality and more attention to the professional development of staff. As to be expected, most effects were found in the HiSE schools, especially the perceived effects on consultation on school functioning and quality (3.0), the achievement orientation (2.4), pupil care (2.2), didactic methods (2.2) and professional development (2.2).

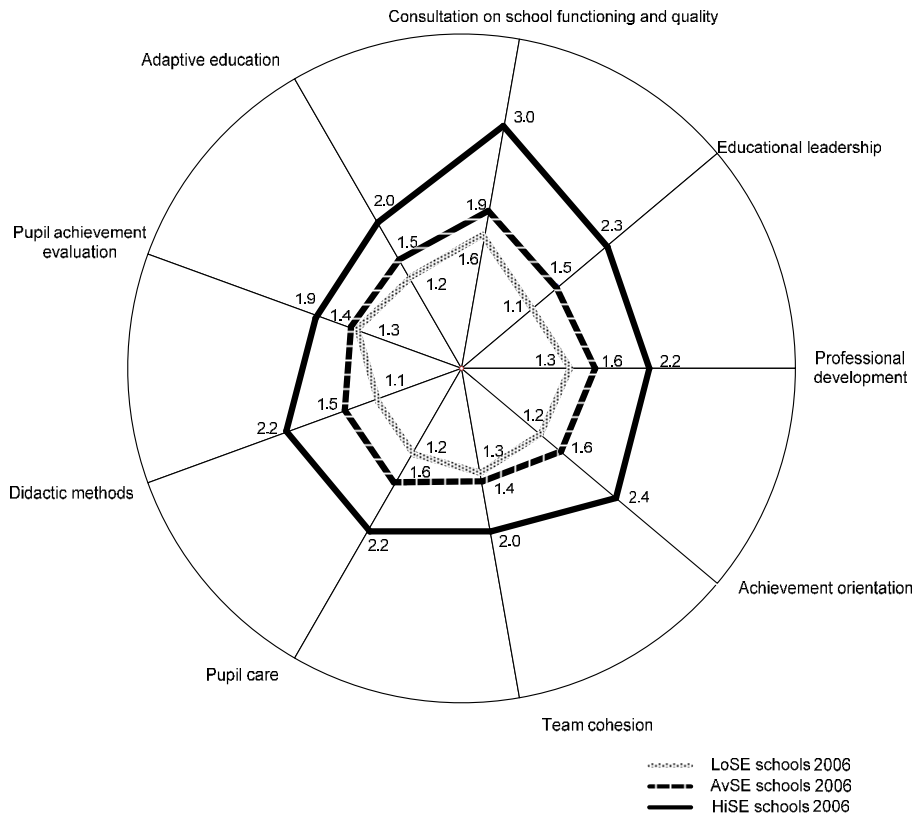


Figure 5.3 Effects of ZEBU use in 2006 in LoSE, AvSE, and HiSE schools. The variables were measured on a four-point scale: 1 = to a minimal degree/not, and 4 = to a great degree)

5.5 A Comparison of the Perceived Effects of ZEBU Use on the Process Indicators

Although ZEBU use had no effect on pupil achievement, it had some minor perceived effects on several process indicators. Table 5.7 shows the main results of the comparison of the perceived effects of ZEBU for 2003, 2004 and 2006 (post hoc ANOVA analyses). The significant increases in perceived effects are printed in bold.

Table 5.7 The comparison of the effects of ZEBO use over the years.

Variables	Time (I)	N (I)	Time (J)	N (J)	Mean difference (J-I)	Std. error
Consultation on school functioning and quality according to principals	2003	40	2004	46	0.10	0.19
	2004	46	2006	25	0.32	0.22
	2003	40	2006	25	0.42	0.23
Consultation on school functioning and quality according to teachers	2003	186	2004	220	0.16	0.08
	2004	220	2006	130	0.20	0.09
	2003	186	2006	130	0.36**	0.10
Educational leadership according to principals	2003	37	2004	41	-0.10	0.17
	2004	41	2006	21	0.17	0.20
	2003	37	2006	21	0.07	0.20
Educational leadership according to teachers	2003	146	2004	158	0.14	0.08
	2004	158	2006	96	0.03	0.10
	2003	146	2006	96	0.18	0.10
Professional development according to principals	2003	39	2004	46	-0.13	0.18
	2004	46	2006	25	0.11	0.21
	2003	39	2006	25	-0.01	0.22
Professional development according to teachers	2003	169	2004	194	0.11	0.08
	2004	194	2006	118	0.19	0.08
	2003	169	2006	118	0.31**	0.09
Achievement orientation according to principals	2003	39	2004	45	0.06	0.19
	2004	45	2006	23	0.18	0.22
	2003	39	2006	23	0.24	0.23
Achievement orientation according to teachers	2003	172	2004	196	-0.07	0.08
	2004	196	2006	117	0.26**	0.09
	2003	172	2006	117	0.18	0.10
Team cohesion according to principals	2003	38	2004	38	0.08	0.17
	2004	38	2006	23	-0.02	0.20
	2003	38	2006	23	0.06	0.20
Team cohesion according to teachers	2003	179	2004	195	0.02	0.07
	2004	195	2006	116	0.11	0.09
	2003	179	2006	116	0.13	0.09
Pupil care according to principals	2003	40	2004	44	0.19	0.20
	2004	44	2006	23	-0.06	0.24
	2003	40	2006	23	0.13	0.24
Pupil care according to teachers	2003	168	2004	194	0.03	0.09
	2004	194	2006	118	0.07	0.10
	2003	168	2006	118	0.10	0.11
Didactic methods according to principals	2003	30	2004	45	0.20	0.16
	2004	45	2006	19	0.26	0.18
	2003	30	2006	19	0.46	0.20
Didactic methods according to teachers	2003	174	2004	202	0.06	0.05
	2004	202	2006	125	0.24**	0.08
	2003	174	2006	125	0.30**	0.08
Pupil achievement evaluation according principals	2003	36	2004	43	0.23	0.18
	2004	43	2006	23	-0.09	0.20
	2003	36	2006	23	0.14	0.21
Pupil achievement evaluation according to teachers	2003	194	2004	207	0.01	0.07
	2004	207	2006	131	0.15	0.08
	2003	194	2006	131	0.16	0.08
Adaptive education according to principals	2003	39	2004	41	0.05	0.17
	2004	41	2006	24	0.12	0.19
	2003	39	2006	24	0.17	0.19
Adaptive education according to teachers	2003	182	2004	197	-0.06	0.08
	2004	197	2006	124	0.18	0.09
	2003	182	2006	124	0.12	0.09

Note: **p < .01.

Table 5.7 shows that the impact of ZEBO on the process indicators as perceived by principals hardly increased over the years. In 2004 and 2006 teachers perceived a significantly stronger effect of ZEBO on four process indicators as compared with 2003 (when the significance was tested at the .01 level). ZEBO use had a growing effect on consultation on school functioning and quality, professional development, achievement orientation, and the didactic methods used by teachers in the classroom. As was evident from the results presented in Chapter 4, teachers started gradually making more use of the ZEBO output over the years, which may explain the increase in the perceived effects of ZEBO on the process indicators, although the effects remain limited.

Principals were found to perceive stronger effects of ZEBO than teachers. This is in line with findings on ZEBO use as reported in Chapter 4. Principals also reported more use of the ZEBO output as compared with teachers. However, the use of the ZEBO output by principals did not increase over time; neither did the effects of ZEBO, as perceived by principals.

Teachers are just starting to use the results. This is also visible in the perceived effects on the process indicators they report. In 2003, for example, only 6% of the teachers (n=220) mentioned that ZEBO use had an effect on the professional development of school staff, in 2006 this percentage had increased to 17% (n=141). This increase is significant ($p < .01$)

5.6 Perceived Effects of ZEBO use on Process Indicators: Conclusion

From the Evaluation of ZEBO Questionnaire results, it became clear that the perceived effects of ZEBO use on the process indicators were limited so far. On several process variables no effect of ZEBO use was found. However, it must be taken into account that the use of the ZEBO output was also limited.

In general, ZEBO use did not have unintended negative effects either. In addition, according to half of the principals and more than one third of the teachers, ZEBO use led to an increased attention to the quality of education. A limited number of schools used the ZEBO output intensively and in these schools some effects on important prerequisites for school improvement were found. ZEBO use in these schools had a positive impact on dialogue and reflection. ZEBO led to an improvement in consultation and communication within these schools. Furthermore, principals and teachers reported an increase in the achievement orientation, an improvement in adaptive education, an improved functioning of the principal, and an increase in professional development activities. Moreover, some teachers indicated that they improved their didactic behaviour as a result of ZEBO use.

Chapter 6

Factors Influencing the Use of ZEBO

6.1 Introduction

This chapter focuses on the factors which have a relationship with the use of ZEBO. In the research framework for this study three groups of factors were hypothesized to influence the use of a School Performance Feedback System such as ZEBO: (A) the characteristics of ZEBO, such as the relevance of the ZEBO output, (B) implementation process features, for example, the clarity of the goal of ZEBO use, and (C) school organisational characteristics of the school into which ZEBO is implemented, such as time and resources available for innovation activities.

To study which factors influence the use of ZEBO, the Evaluation of ZEBO Questionnaire was used. Principals and teachers completed this questionnaire in 2003, 2004, and 2006, each time after schools administered ZEBO. The variables in the framework (see Figure 6.1) were measured by a four point scale, items ranging from “strongly disagree” (1) to “strongly agree” (4).

One additional *school organisational characteristic (C)* which was also expected to influence the use of the ZEBO output, and which was not measured by the questionnaire, is the ZEBO score. If a school scores on or above average on the process indicators measured by ZEBO, there is probably no need, or a reduced need to use the ZEBO output, since this means that the school scored on or above average on the process indicators measured by ZEBO in comparison with schools from a national sample. To take this into account a variable representing the number of scales on which the school scored on, or above average (the ZEBO score) was calculated.

To determine which factors influence the use of ZEBO, correlations were calculated between ZEBO use (D), and the characteristics of ZEBO (A), as perceived by its users, the implementation process features (B), as perceived by the ZEBO users, and the school organisational characteristics (C), as perceived by the ZEBO users. Variables correlating significantly ($p < .01$) with the use of ZEBO were subsequently entered as potential predictors into stepwise regression analyses (in order of size) on ZEBO use, for the data from the principals. The results are presented in section 6.2. For the data from the teachers, multilevel analyses (based on the correlational analyses) were computed, since more than one teacher from every school completed the Evaluation of ZEBO Questionnaire (teachers are nested within schools). Section 6.3 presents the results of these analyses.

It was decided not to compute multiple regression analyses or multilevel analyses at the level of the school as a whole (principals and teachers taken together), since variance analyses showed that teachers and principals differed significantly in their opinion on (A) the characteristics of ZEB0, (B) the implementation process features, and (C) the school organisational characteristics.

Interviews were conducted to further explore why principals and teachers from some schools used the results to make improvements, whereas other principals and teachers did not (see section 6.4).

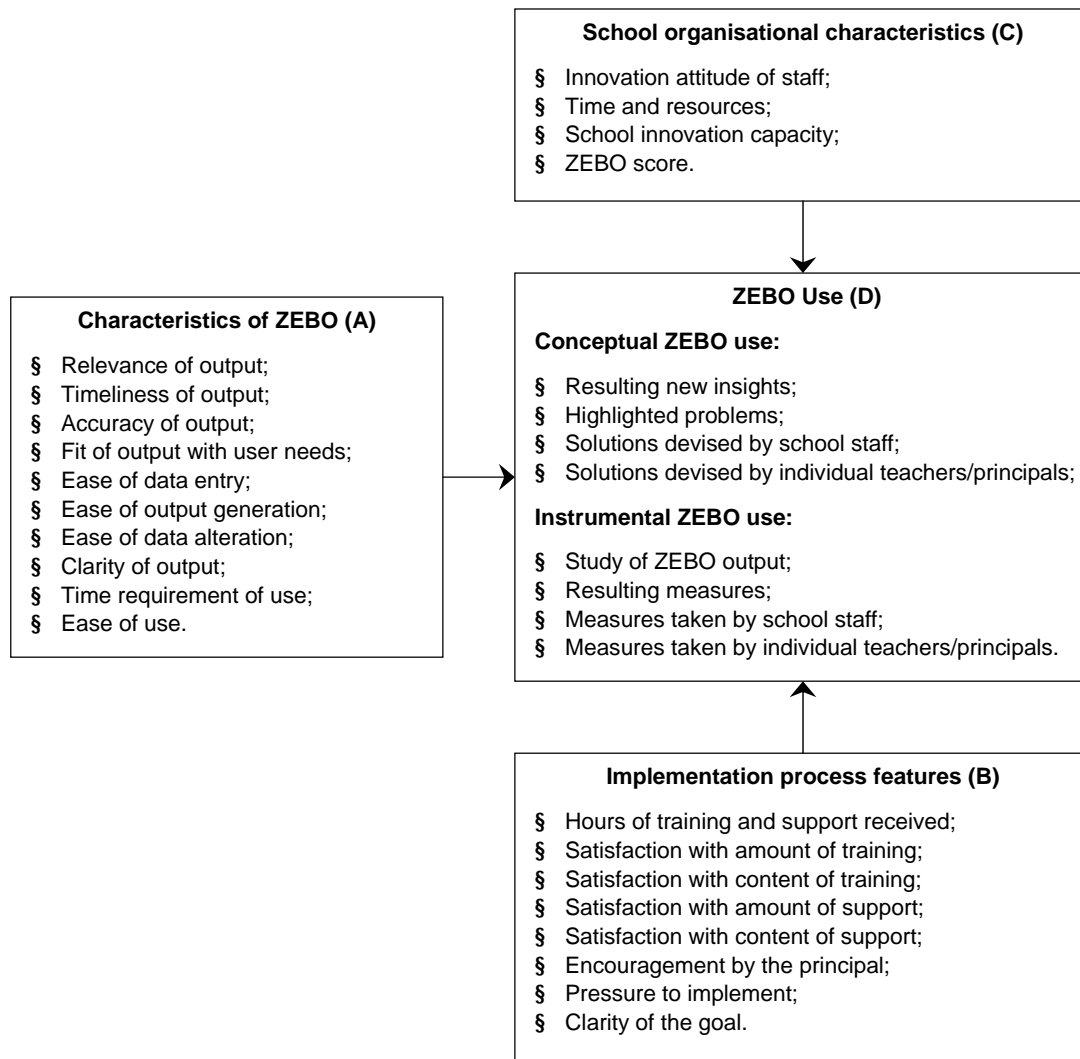


Figure 6.1 Variables hypothesised to influence the use of ZEB0

6.2 Factors Influencing the Use of ZEB0: Results of the Principals

6.2.1 Analysis of Correlations on the Data from the Principals

The assumed relationships between the variables in Figure 6.1 were examined using correlational and multiple regression analyses in 2003, 2004, and 2006. Correlations were calculated between the degree of conceptual ZEB0 use and instrumental ZEB0

use (D), and characteristics of ZEBO (A), implementation process features (B), and school organisational characteristics (C). The significant ($p < .01$) results of the analysis of correlations are presented in Table 6.1. Since both correlational analyses and regression analyses were conducted on the same data set, the alpha level was adjusted downward to 0.01 instead of 0.05 to consider chance capitalisation.

Two (of ten) characteristics of ZEBO (A) were found to correlate positively with the use of ZEBO by principals and the correlations point in the expected direction. The use of ZEBO correlates positively and significantly with the degree to which principals think that the ZEBO output is relevant and the degree to which they feel that the ZEBO output fits with their needs.

Two (of eight) implementation process features (B) were found to correlate positively with the use of ZEBO: the degree to which the goal of using ZEBO is clear, and the degree of encouragement of the use of ZEBO by the principal.

Six (of twenty) characteristics of the school organisation (C) were found to correlate significantly and positively with the use of ZEBO, i.e. the degree to which principals think that:

- § ZEBO use will improve the quality of the school;
- § They are not afraid a lot of things will change because of ZEBO use;
- § As a team, it was decided to participate in the ZEBO project;
- § The school monitors the quality of education;
- § They encourage the professional development of teachers;
- § The school experiments regularly with how to improve education.

Table 6.1 Significant correlations ($p < .01$) between conceptual ZEBO use and instrumental use, and characteristics of ZEBO (A), implementation process features (B), and school organisational characteristics (C), as perceived by principals

	Conceptual use				Instrumental use							
	2003	N ^a	2004	N ^a	2006	N ^a	2003	N ^a	2004	N ^a	2006	N ^a
<i>Characteristics of ZEBO</i>												
Relevance of output	.43	39							.48	38		
Fit of output with user needs	.45	38										
<i>Implementation process features</i>												
Clarity of goal					.53	22					.71	20
Encouragement by principal									.42	38	.64	20
<i>School organisational characteristics</i>												
ZEBO leads to quality improvement			.35	40	.55	24					.54	22
Not afraid of changes			.46	38								
Team decision											.72	17
Monitor quality of education							.37	34				
Principal encouragement of professional development									.47	39		
Experiment to improve education							.36	34				

Note: ^aNumber of principals who filled in all the questions on the use of ZEBO.

Correlations do not give information on the predictive power of the variables. Multiple regression analyses may give a better understanding of how the characteristics of ZEBO (A), implementation process features (B), and the school organisational characteristics (C) are predictive for the use of ZEBO (D). To explore the relationships between the variables in the theoretical framework a series of multiple regression analyses was carried out for the principals. The results are presented in the next section.

6.2.2 Multiple Regression Analyses Based on the Data from the Principals

The variables which correlate significantly with the conceptual use of ZEBO were entered into stepwise regression analyses with the conceptual use of ZEBO as the dependent variable, for the data collected in 2003, 2004 and 2006. The variables which correlate significantly with instrumental use were entered into stepwise regression analyses on the instrumental use. The number of available cases (N=41 in 2003, N=48 in 2004, and N=25 in 2006) required a selection of potential predictors. A maximum of 4 variables (2 in 2006) was entered in each regression analysis: those variables which proved to correlate most strongly and significantly with the use of ZEBO. It is recognised that this may result in chance capitalisation. However, considering the number of variables in the theoretical framework, which are all expected to influence the use of ZEBO, it was necessary to come to a reduction of variables (to be entered into the regression analyses). Since it was not possible to come to a theoretical reduction because all variables seemed equally important, it was decided to make a reduction based on the correlational analyses. Table 6.2 displays the results of the analyses.

In 2003, three of the four variables entered were found to explain variance in the use of ZEBO at the level of the principal. One ZEBO characteristic (A) was found to influence the use of ZEBO: the extent to which the ZEBO output fits with the needs of users. Two school organisational factors (C) were found to influence the use of ZEBO: the degree to which the school monitors the quality of its functioning, and the degree to which school staff experiments to improve education. These variables together explain 20% of the variance in the conceptual use of ZEBO, and 27% of the variance in the degree of instrumental ZEBO use.

In 2004, four of five variables explained the variance in the use of ZEBO (31% of the variance in conceptual use, and 33% of the variance in instrumental use), as perceived by principals. Again, the ZEBO characteristic (A), the extent to which the ZEBO output fits with the needs of the users, was found to influence the use of ZEBO. The degree to which principals think the use of ZEBO will lead to quality improvement, the degree to which the principal is not afraid of changes because of ZEBO, and the degree to which the principal encourages the professional development of school staff are school organisational factors (C) which were found to have influenced the use of ZEBO.

Three of the six variables explained variance in the use of ZEBO (D) in 2006. One school organisational factor (C), the degree to which principals think the use of ZEBO will lead to quality improvement, explains 30% of the variance in conceptual ZEBO use. One implementation process factor (B), the degree to which the goal of ZEBO is clear, and one school organisational factor (C), the degree to which the decision to participate in the ZEBO project was made by the team explained 69% of the variance in instrumental use.

Table 6.2 Significant regression weights ($p < .05$) for regression analyses predicting conceptual use and instrumental use based on the principal data for 2003, 2004, and 2006

Independent variable	Coefficients	Conceptual use			Instrumental use		
		2003 ^a	2004 ^b	2006 ^c	2003 ^d	2004 ^e	2006 ^f
Fit of output with user needs	B	1.59			1.47		
	β	0.45			0.27		
	Sig.	0.00			0.03		
Clarity of goal	B						2.93
	β						0.47
	Sig.						0.01
ZEBO leads to quality improvement	B		1.25	1.98			
	β		0.31	0.55			
	Sig.		0.03	0.01			
Not afraid of changes	B		2.12				
	β		0.44				
	Sig.		0.00				
Team decision	B						1.75
	β						0.50
	Sig.						0.01
Principal encouragement of professional development	B				2.50		
	β				0.35		
	Sig.				0.04		
Monitor the quality of education	B				1.97		
	β				0.38		
	Sig.				0.01		
Experiment to improve education	B				1.54		
	β				0.37		
	Sig.				0.02		
	R ²	0.20	0.31	0.30	0.27	0.33	0.69
	Sig.	0.00	0.00	0.01	0.00	0.00	0.00

Note. ^an = 39; ^bn = 38; ^cn = 21; ^dn = 34; ^en = 32; ^fn = 32.

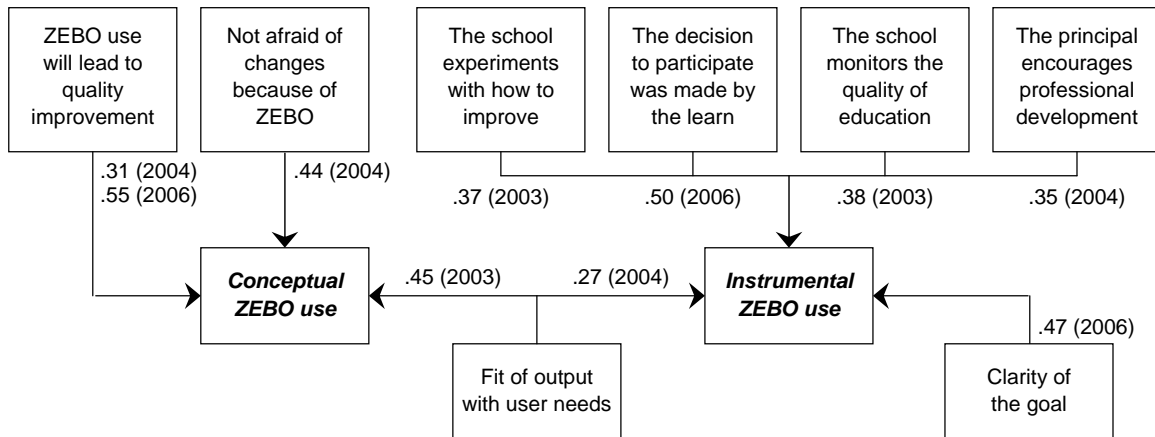


Figure 6.2 Standardised regression coefficients of significant relationships for the conceptual and instrumental use of ZEBU, by principals in 2003, 2004, and 2006

All significant coefficients of the variables influencing the use of ZEBU by principals are shown in Figure 6.2. Contrary to what was expected (based on the analysis of correlations presented before) the two other variables (the relevance of the ZEBU output, and the encouragement of the use of ZEBU by the principal) entered, explained no variance additional to the variables in Figure 6.2. First it was theorised that this was caused by multicollinearity. If a strong correlation (0.90 and above) exists between two or more predictors in a regression model, this is called multicollinearity. High levels of collinearity increase the probability that a good predictor of the dependent variable will be found non-significant and rejected from the model (Field, 2000).

To test for multicollinearity, collinearity diagnostics were computed using SPSS, by calculating the Variance Inflation Factor (VIF). The VIF indicates whether a predictor has a strong linear relationship with other predictors, or not (Field, 2000). For our current models the Variance Inflation Factor (VIF) values are well below 10. VIF values of 10 or more should be a cause for concern (Bowerman & O'Connell, 1990 in Field 2000). Therefore, it may be concluded that collinearity does not exist within the data entered in the regression analysis, although some of the independent variables are correlated with each other. The independent variables which are predictive for ZEBU use probably predict part of the same variance as those independent variables which (1) correlate with the use of ZEBU, (2) which were entered in the regression analyses, but (3) which were excluded from the regression model by SPSS (see Appendix 6.1).

6.3 Factors Influencing the Use of ZEBU: Results of the Teachers

6.3.1 Analysis of Correlations on the Data from the Teachers

To answer the question which variables (shown in Figure 6.1) have a relationship with the use of ZEBU by teachers, correlations were calculated between (A) the characteristics of ZEBU as perceived by teachers, (B) the implementation process features, and (C) school organisational characteristics as perceived by teachers, and

(instrumental and/or conceptual) ZEBO use (D). Variables which correlate significantly with the use of ZEBO were then entered as potential predictors into multilevel analyses. Because the teacher data collected in this study have a nested structure (teachers are nested within schools), multilevel analyses were required (for further details see Chapter 3). The results of the analyses may be found below. First, the results of the correlation analyses (see Table 6.3) are discussed in this section. In the next subsection the results of the multilevel regression analyses are discussed.

The following four (of ten) characteristics of ZEBO (A) were found to correlate significantly ($p < .01$) with the use of ZEBO by teachers (D) and point in the expected direction (positive correlations), i.e. the degree to which teachers think that:

- § The ZEBO output fits with the needs of the school;
- § It is easy to enter data into ZEBO;
- § The ZEBO output is accurate;
- § Working with ZEBO does not take a lot of time.

Several (five of eight) implementation process features (B) were also found to correlate positively and significantly with the use of ZEBO (D). The use of ZEBO correlated positively and significantly with:

- § The amount of training and support received;
- § The degree to which teachers feel they received sufficient support;
- § The degree to which teachers feel they received sufficient training;
- § The degree of encouragement to use ZEBO by the principal;
- § The degree of clarity of the goal of ZEBO.

Eight (of the twenty) school organisational characteristics (C) were found to correlate significantly and positively with the conceptual use of ZEBO (D), i.e. the degree to which teachers think that:

- § The use of ZEBO will improve the quality of our school;
- § They are not afraid that a lot of things change because of ZEBO use;
- § The school reserved extra time and resources for the use of ZEBO;
- § The decision to participate in the ZEBO project was taken by the team;
- § They can influence the measures taken as a result of the ZEBO output;
- § The school monitors the quality of education;
- § The principal encourages professional development;
- § They exchange information on their functioning.

One characteristic of the school organisation (C) correlated significantly with the use of ZEBO (D), but did not point in the expected direction. The use of ZEBO correlates negatively with the degree to which strong team cohesion exists. The reason for this negative correlation is not clear. One would expect that strong team cohesion is a condition required for, or at least promoting, the use of self-evaluation results. One

possible explanation is that schools with less strong team cohesion have a greater need for use of an instrument such as ZEBO.

One school organisational characteristic (C), the ZEBO score (the number of ZEBO scales on which the school scored on or above average) correlated significantly and negatively with the use of ZEBO (D). As expected, relatively low average ZEBO scores (i.e. the school scored below average on several ZEBO scales) may promote the use of the ZEBO output by teachers.

Table 6.3 Significant correlations ($p < .01$) between conceptual ZEBO use and the instrumental ZEBO use and the characteristics of ZEBO (A), implementation process features (B), and the school organisational characteristics (C) for the teacher data

	Conceptual use				Instrumental use							
	2003	n^a	2004	n^a	2003	n^a	2004	n^a	2006	n^a		
<i>Characteristics of ZEBO (A)</i>												
Fit of output with user needs	.43	120	.26	133			.37	85	.25	80	.29	71
Ease of data entry			.32	147			.26	68	.31	104		
Accuracy of output									.32	80		
Time requirement of use									.31	104		
<i>Implementation process features (B)</i>												
Hours of training and support received									.27	104		
Satisfaction with amount of training			.31	147					.27	104		
Satisfaction with amount of support			.26	96					.26	65		
Encouragement by principal			.30	147			.41	85	.35	98		
Clarity of goal	.44	120					.47	85	.41	104		
<i>School organisational characteristics (C)</i>												
ZEBO leads to quality improvement	.42	120	.33	136			.41	85			.27	72
Not afraid of changes			.30	147								
Time and resources			.28	147	.41	75					.44	73
Team decision	.22	120							.29	84		
Teachers influence ZEBO measures			.28	147			.29	85	.29	104		
Monitor the quality of education									.34	100		
Principal encouragement of professional development									.27	104		
Teachers exchange information											.26	72
Team cohesion					-.28	100						
ZEBO score	-.31	120					-.31	111				

Note: ^aNumber of teachers who filled in all the questions regarding the use of ZEBO.

6.3.2 Multilevel Analyses Based on the Data from the Teachers

The five variables which correlated significantly (positively or negatively) with the conceptual use of ZEBO in 2003 were entered into multilevel analyses with the conceptual use of ZEBO as the dependent variable. Two models are presented in Table 6.4: Model 0, which is the basic model without the explaining independent variables, and Model 1. Model 1 differs from Model 0 in that the implementation process and school organisational variables, which explained variance in the conceptual use of ZEBO, have been added. Model 1 fits the data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2=37$, $df=3$, $p=0.00$) was significant. The implementation process feature (B), the clarity of the goal of ZEBO, and the two school organisational characteristics (C), i.e. the degree to which teachers thought that the use of ZEBO would lead to quality improvement, and the degree to which the decision to participate in the ZEBO-project was made by the team, explained 29% of the variance between schools and 27% of the variance between teachers in the conceptual use of ZEBO in 2003.

Table 6.4 Variables influencing the conceptual use of ZEBO (2003)

	Model 0 (N=120)		Model 1 (N=120)	
	Est.	S.E.	Est.	S.E.
<i>Fixed effects</i>				
Intercept	-0.02	0.12	-0.04	0.10
clarity of goal of ZEBO			0.18	0.08
ZEBO leads to quality improvement			0.36	0.09
Team decision			0.15	0.07
<i>Variance components</i>				
Between schools	0.42	0.13	0.29	0.09
Between teachers	0.33	0.05	0.26	0.04
<i>Percentage explained</i>				
Between schools			29	
Between teachers			27	
<i>Deviance</i>				
Improvement in model fit (p)	267		230	0.00

The seven variables which correlated significantly with the instrumental use in 2003 were entered into multilevel analyses on the instrumental use of ZEBO (Table 6.5). Model 1, with all the significant variables for instrumental use, fits the data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2=28$, $df=4$, $p=0.00$) was significant. Two implementation process features (B), namely the degree of encouragement to use ZEBO by the principal and the clarity of the goal of ZEBO, and two school organisational characteristics (C), i.e. the degree to which teachers feel they can influence measures taken based on ZEBO and the ZEBO score (number of ZEBO scales on which the school scored on or above average), together explained 44% of the variance between schools and 38% of the variance between teachers in the instrumental use of ZEBO (D).

Table 6.5 Variables influencing the instrumental use of ZEBO (2003)

	Model 0 (N=85)		Model 1 (N=85)	
	Est.	S.E.	Est.	S.E.
<i>Fixed effects</i>				
Intercept	0.08	0.15	-0.01	0.11
Encouragement by principal			0.26	0.09
Clarity of goal			0.23	0.09
Teachers influence on ZEBO measures			0.22	0.09
ZEBO score			-0.27	0.12
<i>Variance components</i>				
Between schools	0.70	0.19	0.36	0.11
Between teachers	0.23	0.05	0.22	0.05
<i>Percentage explained</i>				
Between schools			44	
Between teachers			38	
<i>Deviance</i>				
Improvement in model fit (p)	189		161	0.00

Multilevel analyses on the 2004 and 2006 teacher data were conducted in the same manner. The detailed results may be found in Appendices 6.2 and 6.3.

In 2004, one of the ZEBO characteristics (A) explained variance in the use of ZEBO (D), i.e. the degree to which it is easy to enter data. Two implementation process features (B), i.e. the degree of satisfaction with the amount of training and the degree to which the principal encouraged the use of ZEBO, explained variance in the use of ZEBO (D), and three school organisational characteristics (C), i.e. the extent to which time and resources are available, the degree to which teachers are not afraid of changes because of the use of ZEBO, and the degree to which teachers feel that they may influence measures taken based on the ZEBO output, explained variance in the use of ZEBO (D). Together, these six variables explained 27% of the variance between schools and 31% of the variance between teachers in the conceptual use of ZEBO. The improvement fit for Model 1 for conceptual ZEBO use was significant ($\chi^2=50$, $df=6$, $p=0.00$).

One ZEBO characteristic (A), the amount of time working with ZEBO costs; one implementation process feature (B), the degree of encouragement of the use of ZEBO by the principal; and two school organisational characteristics (C), the degree to which teachers feel they can influence measures taken based on ZEBO and the degree to which the principal encourages professional development, explained 30% of the variance between schools and 30% of the variance between teachers in the instrumental use of ZEBO (D) in 2004. The improvement fit for this model was also significant ($\chi^2=40$, $df=4$, $p=0.00$).

In 2006, two school organisational characteristics (C), i.e. the degree to which the school has time and resources available for the use of ZEBO and the degree of team cohesion, explained variance in the conceptual use of ZEBO (D) (25% between schools and 42% between teachers). The model with these variables fit the data significantly better than the basic model and the improvement fit ($\chi^2=17$, $df=2$, $p=0.00$) was significant.

Model 1 for instrumental ZEB0 use also fits the data significantly better than the basic model, and the improvement fit for Model 1 ($\chi^2=23$, $df=3$, $p=0.00$) was significant. Three school organisational characteristics (C), i.e. the amount of time and resources the school has available for the use of ZEB0, the degree to which teachers exchange information on their functioning and the degree to which working with ZEB0 will lead to quality improvement, according to the teachers, explained variance (29% between schools and 30% between teachers) in the instrumental use of ZEB0 (D) in 2006. The combined 2003, 2004, and 2006 results may be found in Figure 6.3.

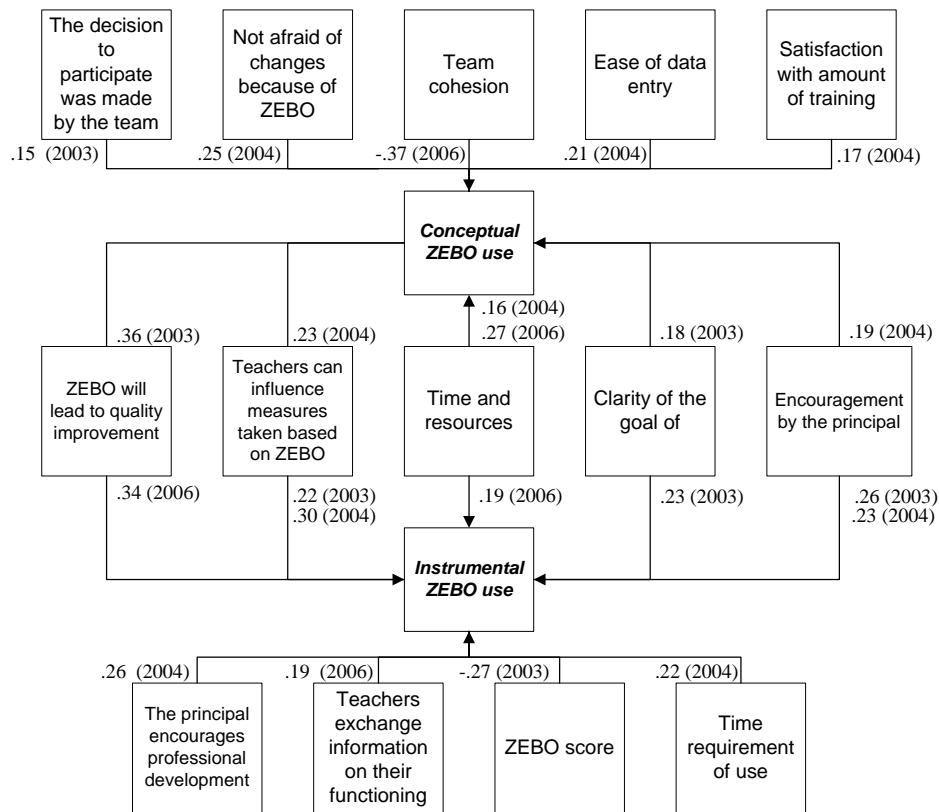


Figure 6.3 Standardised regression coefficients of significant relationships for the conceptual and instrumental use of ZEB0 by teachers in 2003, 2004, and 2006

The other five variables entered explained no variance additional to the 14 variables in Figure 6.3. Collinearity diagnostics were computed in SPSS. Based on these diagnostics it was concluded that collinearity does not exist between the predictors within the teacher data. The variables which are predictive for ZEB0 use probably predict part of the same variance as those independent variables which correlate with the use of ZEB0, but which were excluded from the multilevel models (see Appendix 6.4).

6.4 The Results of the Interviews in 2003, 2004, and 2006

Based on the Evaluation of ZEBO Questionnaire results in 2003, 2004 and 2006, the ZEBO schools were divided into three groups (for further details see Chapter 3):

- § One group of schools which did not use the ZEBO output (Low Self Evaluation, “LoSE”);
- § One group of schools which made average use of the ZEBO output (Average Self Evaluation, “AvSE”);
- § One group of schools which used the ZEBO output intensively (High Self Evaluation, “HiSE”).

Respondents from LoSE, AvSE, and HiSE schools were interviewed to validate the Evaluation of ZEBO Questionnaire results and to further explore the factors promoting ZEBO use. In 2003, thirty-one respondents were interviewed. In 2005 (using the 2004 questionnaire as a starting point), 25 respondents were interviewed. In 2006, due to time restrictions, it was not possible to conduct interviews with respondents individually. As an alternative, a focus group interview was set up. However, in November when the focus group was planned to take place, only two of the ten invited participants were able to join. Two principals (one each from a HiSE school and an AvSE school) were interviewed simultaneously. The major results are presented below. Detailed results may be found in Appendices 6.5, Appendix 6.6, and Appendix 6.7.

The 2003, 2004, and 2006 interview results also indicate that implementation process features and school organisational characteristics play an important role in the use of the ZEBO output. Characteristics of ZEBO seem to be less important, since all respondents judge these characteristics more or less positively.

Firstly, the interview results confirm the Evaluation of ZEBO Questionnaire results, in that the principal plays an important role in the use of self-evaluation results, especially at the start of the school self-evaluation process. The principals from the LoSE schools (and to some extent the principals from the AvSE schools) were less active in encouraging the use of ZEBO, and in encouraging the professional development of teachers. Principals from HiSE schools explained repeatedly to their staff why the school was using ZEBO, what the benefits were for the school, and how the school was going to use the results. One principal put it as follows:

I encourage the use of ZEBO by giving ZEBO a lot of attention in team meetings and by trying to explain the benefits of it. In education, things cannot take up too much time and energy. People say that these things add to their workload, which is already high. I have tried to make the teachers see that they should not perceive it as something extra they have to do, but that they have to see it as a learning process for the school. In general, teachers were positive

about this. Furthermore, we discussed the results in team meetings: we shared the results, and discussed our opinions on these results. By giving it a lot of attention the enthusiasm lasts.

Principals sometimes must persuade the team to engage in a self-evaluation process. One of the principals stated, for example, in 2006:

I had to persuade the team to participate. The first reaction of the team was “there she is again with one of her things”. I had to explain to them why it is important to start with self-evaluation. After working with ZEBO they are more positive. They are even curious about the results.

Another difference between the LoSE and AvSE schools and the HiSE schools may be found in the requirement for training and support. None of the respondents from the HiSE schools indicated that they could have used training and support in the use of ZEBO. Respondents from LoSE and AvSE schools responded that they could have used training and support. They needed training in helping pupils completing the questionnaires, in interpreting the questions, in analysing the results, and support in integrating ZEBO into the school policy. A teacher from an AvSE school (in 2005) answered to the question what kind of support the school required in the use of ZEBO with:

[Training and support] in a team meeting, in which an external person would explain to us what the goal of using ZEBO is, and how to complete the questionnaire. The information currently is minimal. I think that if an expert gives a clear presentation before we use it, and when teachers can ask questions, people will be much more motivated. Currently, it is just another thing we have to do in combination with our other activities, and that is a shame. I think it is a very useful instrument, but attention has to be given to it. And this is not the case.

By 2004 and again in 2006, all respondents were clear about the goal of ZEBO. However, in 2003, three teachers and one principal from LoSE schools were not familiar with the goal of ZEBO. They mentioned a wide range of (sometimes very unclear) goals, such as participating in research and using an instrument because they were obliged to, by law. When asked to elaborate on these goals, the teachers and principal were not able to explain this further. One of the teachers stated that she had no idea what the goal of ZEBO was:

I have no idea; it is just an instrument that we are obliged to use that came from your university.

Another difference between the LoSE, AvSE, and HiSE schools, found in 2003, relates to the school organisation and concerns the use of other self-evaluation and quality care instruments. Four teachers from two of the schools which did not use the ZEBO output, indicated that they were also using another instrument. These schools stopped participating in the ZEBO project in 2004.

It is also possible that principals and teachers are doing more with the ZEBO output than some teachers think. One of the principals stated in 2006 that her school is using the ZEBO output to make improvements and teachers are working towards these improvements. The school implemented classroom consultation but the principal doubted that most teachers knew this was a direct result of ZEBO.

Furthermore, respondents from LoSE schools expressed in 2003, 2004, and 2006 a more negative attitude against ZEBO than respondents from AvSE and HiSE schools, such as it is “another we are obliged to do”, “another questionnaire to fill out”, and “it only tells me things I already know”.

Respondents from LoSE schools also may feel less ownership over the results. Three principals and six teachers from LoSE schools (in 2003) indicated, for example, that the decision to use ZEBO was taken by their principal or board of governors.

In addition, respondents mentioned in the interviews, three other possible explanations, not measured by the Evaluation of ZEBO Questionnaire, for the limited use of ZEBO in some schools. Firstly, an additional possible explanation for why the LoSE schools did not use the results may be found in the interpretation of the results by the principals. In 2005, in one of the schools, school staff had judged the principal very negatively, and it was reported that severe communication problems existed. The principal studied the results and came to the conclusion that the teachers interpreted the questions wrongly and therefore he considered the results as an inaccurate picture of the school.

Secondly, in the HiSE and AvSE schools there was usually one person (usually the principal, but sometimes the internal educational advisor) who took the responsibility of studying the results and who decided which aspects should be discussed and/or improved. The following quote illustrates this:

I structured the ZEBO-output: I wrote down all the sub-items on which we scored below average. In several meetings we discussed the points we were satisfied with, and which topics are issues of concern and therefore needed improvement.

Thirdly, the interview results further show that an open atmosphere in which everything can be discussed may also promote the use of self-evaluation results. One of the principals put it as follows in 2006:

Sometimes the results are very critical. Some schools might not have the courage to “open the can of worms”. In these schools conflict will arise and people will start talking at cross-purposes instead of talking with each other. For self-evaluation an open atmosphere is needed. It must be possible to discuss issues.

6.5 Conclusions

Both quantitative (questionnaires) and qualitative (interviews) data were used to assess the explanatory power of the theoretical framework. This framework includes three groups of factors potentially influencing the use of the ZEBO output by schools. Table 6.6 shows the results of all three evaluations regarding the variables which were found to influence the use of ZEBO. Teachers and principals differ to some extent with regard to the variables influencing the use of ZEBO. One difference between teachers and principals is that the use of ZEBO by teachers is influenced by implementation process features, whereas this is only minimally so for the principals. A possible explanation for this may be that principals in this study were more involved in the decisions regarding the adoption and implementation of ZEBO. This demonstrates the importance of involving teachers in both the implementation and the use of a SPFS.

Furthermore, the degree to which time and resources were available for the use of ZEBO influenced the use of ZEBO by teachers, but the degree to which time and resources were available did not influence ZEBO use by principals. Principals had to fill out only one ZEBO questionnaire (from the ZEBO instrument) in 2003, 2004, and 2006. Teachers also had to fill out one questionnaire but they were in most cases also responsible for seeing to it that pupils completed the questionnaires. Therefore, they may have felt the need for extra time and resources, in terms of organising the whole ZEBO administration.

Table 6.6 Factors influencing the use of ZEB0 in 2003, 2004 and 2006

ZEB0 use 2003	ZEB0 use 2004	ZEB0 use 2006
<i>ZEB0 characteristics as perceived by its users (A)</i>		
Fit of output with user needs (PQ & I)	Fit of output with user needs (PQ)	
	Ease of data entry (TQ)	
	Time requirements of use (TQ)	
<i>Implementation process features (B)</i>		
Clarity of goal (TQ & I)	Satisfaction with amount of training (TQ & I)	Clarity of goal (PQ)
Encouragement by principal (TQ & I)	Encouragement by principal (TQ & I)	
Somebody in the school takes responsibility over the results (I)	Somebody in the school takes responsibility over the results (I)	
<i>School organisational characteristics (C)</i>		
ZEB0 score (TQ)	Time and resources (TQ)	Time and resources (TQ)
Aspects of the innovation attitude scale:	Aspects of the innovation attitude scale:	Aspects of the innovation attitude scale:
§ ZEB0 leads to quality improvement (TQ)	§ ZEB0 leads to quality improvement (PQ)	§ ZEB0 use leads to quality improvement (PQ & TQ)
§ Teachers influence ZEB0 measures (TQ)	§ Teachers influence ZEB0 measures (TQ)	
	§ Not afraid of changes (PQ & TQ)	
Aspects of the innovation capacity scale:	Aspect of the innovation capacity scale:	Aspects of the innovation capacity scale:
§ Team decision (TQ & I)	§ Principal encouragement professional development (PQ, TQ & I)	§ Team decision (PQ)
§ Monitors the quality of education (PQ)		§ Teachers exchange information (TQ)
§ Experiments to improve education (PQ)	§ Congruent results with the expectations of the principal (I)	§ Team cohesion (TQ)
Congruent results with the expectations of the principal (I)		§ An open climate (I)

Note: PQ: principal questionnaire, TQ: teacher questionnaire, I: interviews. Variables printed in bold are those which were found, more than once, to influence the use of ZEB0.

Variables which were found to influence the use of ZEB0 two or three times (printed in bold in Table 6.6), provide stronger empirical proof for the importance of these variables, in a school self-evaluation process. The ZEB0 evaluation results suggest that several variables, from the three groups of factors (A), (B) and (C) influenced the use of the ZEB0 output at least twice. One characteristic of ZEB0 (A) influenced the use of the results twice:

- § The degree to which the ZEB0 output fits with the needs of the users. This finding is consistent with Grasso's (2003) assumption, that evaluation results should address

the needs of the users, in order to improve the extent to which an evaluation is in fact used, and with Fullan (1991), who states that staff involved in an innovation must feel the need for that innovation. Leithwood et al. (2001) also state that it is important that an information need exists. Potential users should perceive a gap in their knowledge which ZEBO output may help fill.

The results indicate that two implementation process features (B) of the theoretical framework and two additional variables from the interview results, influenced the use of ZEBO twice:

- § The clarity of the goal of ZEBO. The findings suggest that at the start of the school self-evaluation process, the goal of using ZEBO should be clear to all school staff. In 2003, however, at the start of this study, the goal of ZEBO was unclear for several teachers. In 2006, the goal was unclear for several principals, this was probably due to management changes in several schools. Clift et al. (1987) conclude that for the self-evaluation to be successful, it is important that the purpose of the self-evaluation is clear. According to Kyriakides and Campbell (2004) it is important to establish clarity and consensus about the aims of the self-evaluation;
- § The encouragement to use ZEBO provided by the principal. Saunders (2000) concludes, from a study which explored the potential role of value-added analyses in school improvement, that active support of data by the management is positively correlated with schools' utilisation of performance data. Kells (1995) puts it even more strongly. In his view, unless the leader of an organisation (in this case, the principal) is interested and willing to use the self-evaluation results, one should not proceed with the evaluation;
- § The results of the interviews suggest that the ZEBO output must be congruent with the expectations of the principal. This was also found by Cousins and Leithwood (1986) in their review of empirical research conducted on the use of evaluation results;
- § The interview results also suggest that it is also important that the principal (or another person in the school) takes responsibility for the results: distributing the results, explaining them where necessary and discussing with school staff how to use the results to improve the functioning of the school.

Furthermore, four school organisational characteristics (C) influenced the use of ZEBO twice and one variable influenced the use of ZEBO three times, i.e. during all three evaluations:

- § The degree to which teachers and principals feel that the use of ZEBO would lead to quality improvement. This variable is part of the innovation attitude scale and was found to have influenced the use of ZEBO output three times. Cousins and Leithwood (1986) found a positive correlation between evaluation utilisation and users' attitude towards evaluation. The interview results also confirm that the innovation attitude may play a role in the use of self-evaluation findings. A more

negative attitude towards the use of ZEBO was expressed by the teachers interviewed from the LoSE schools, than by those from the AvSE and HiSE schools, especially in 2003. This negative attitude was reinforced by the fact that the schools did not use the results, as a result of which, the ZEBO-administration was a waste of time in the eyes of these teachers. It is difficult to say what is cause and what is effect here: does the negative attitude lead to a less intensive use of ZEBO, or does not using ZEBO fully make users more negative?;

- § The degree to which teachers feel they can influence measures based on the ZEBO output. This variable is also part of the innovation attitude scale. If teachers feel they can influence these measures, this may lead to a feeling of ownership of the results and a sense of empowerment. Davies and Rudd (2001) conclude that it is important to promote ownership of the (results of the) self-evaluation among teachers. Teachers should have the opportunity to focus on aspects of the school that they have identified as areas requiring improvement and take measures accordingly;
- § The degree to which it was felt that the decision to participate was taken by the team, is also likely to promote ZEBO use. This variable is a variable from the innovation capacity scale. Geijsel (2001) defines the innovation capacity of schools as the capacity of schools to implement innovations in a successful manner. The interview results show something similar; in three of four LoSE schools, it was decided by the school board without consultation with the team, to participate in ZEBO. It is only to be expected that these teachers and principals may not feel ownership of ZEBO, and are more negative about ZEBO than a school who felt that they took the decision, as a team to participate in ZEBO;
- § The amount of time and resources the school reserved for the use of ZEBO. 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, play an important role in the use of self-evaluation results;

Several variables were found to influence the use of the ZEBO output only once, suggesting that there is less empirical support for the importance of these variables in the use of self-evaluation results. However, with regard to some of these variables, it is plausible that these variables may only influence the use of the self-evaluation results once. The following is a discussion of some variables which were found to influence ZEBO use, once.

- § Two characteristics of ZEBO (A), namely the degree to which it is easy to enter data in ZEBO and the degree to which working with ZEBO does not take a lot of time. These variables were not found to play a role, any longer, in 2006. A possible explanation for this is that either, schools judged the characteristics positively and chose to continue with ZEBO or they judged the characteristics negatively and chose not to continue. It is also possible that a practice effect occurred. By the third administration data entry may have been easy and quick;

- § The degree to which teachers felt they received sufficient training in the use of ZEBO (an (B) implementation process feature). Davies and Rudd (2001) recommend training for school managers and teachers to equip them to support and carry out their self-evaluations. Providing training and support may help to prevent school self-evaluation being seen as a one-off exercise, rather than as an integral part of school policy. Scheerens et al. (2003) state that, without external support, school self-evaluation is likely to fail. Schools often have problems in connecting “diagnosis” and “therapy”. In the interviews, several principals and teachers indicated that they felt a need for more support in interpreting the ZEBO output and designing remediation and corrective actions to improve the school’s functioning. In 2006, teachers and principals indicated that they did not require any (further) training, for example, because they received training from, for example, the school advisory service. However, the results also indicate that most schools still did not make use of the ZEBO output in 2006, to improve their functioning. Although several principals and teachers indicated that they experienced sufficient training and support in the use of evaluation results, providing them with higher levels of training and support may lead to higher levels of evaluation use. Moreover, they may not have realised the benefits of training;
- § Firstly, various aspects of the innovation capacity scale (a school organisational characteristic (C)), including the degree to which school staff experiments with ways to improve education, the degree to which the school monitors the quality of education, the degree to which teachers exchange information on their functioning, and the degree to which the principal encourages the professional development of school staff. Geijsel (2001), for example, concludes that the support principals provide for the professional development of teachers influences teachers’ changes in practices. These aspects may only have influenced the use of ZEBO once, but this may be due to the lack of variance in these variables during the other evaluations. Closer analyses showed, for example, that unlike in 2003, all principals judged the aspects of the school innovation capacity similarly in 2006. Other aspects of the innovation capacity, as described above, seem consistently important in the use of school self-evaluation results;
- § Another school organisational characteristic (C) is the ZEBO score. In 2003, schools which scored lower on the ZEBO scales were more likely to make more use of the results. This variable did not play a role any longer in 2004 or 2006. More schools scored on or above average on the ZEBO scales in 2004 and 2006, than had in 2003, making the variance between schools smaller;
- § Thirdly, the interview results suggest that another variable may influence the use of self-evaluation results, namely the degree to which an open climate exists in which people are not afraid of confrontation and critique and in which everything is open for discussion. The interviews suggest that this variable may have played a role, but this could not be verified quantitatively, because it was not included in the Evaluation of ZEBO Questionnaire.

Concluding, characteristics of ZEBO (A) only seem to influence the use of the self-evaluation results at the start of a school self-evaluation process. Schools which judged the ZEBO characteristics negatively stopped using ZEBO and continued with another instrument. In 2003 and 2004, for example, 11 schools judged the characteristics negatively and chose not to continue with ZEBO. Seven characteristics of ZEBO, as perceived by its users, contrary to what was expected, did not influence the use of ZEBO output. The fact that several characteristics of ZEBO, as perceived by its users, were judged positively by the majority of respondents (leading to less variance) may explain why several variables were not found to influence the use of the ZEBO output.

Implementation process features (B) also seem to play a role at the start of a school self-evaluation process only. The clarity of the goal, encouragement by the principal, and training all play an important role during the implementation of school self-evaluation, but appear not to later. The findings suggest that in 2006, encouragement of ZEBO use by the principal was no longer required. However, analysis showed that 86% of the teachers responded that the principal still encouraged the use of ZEBO, suggesting that the role of the principal is continuously important.

Implementation process features, which were not found to influence the use of the ZEBO output, relate to the pressure to implement ZEBO and the (amount of and satisfaction with) support. It is possible that these variables separately do not influence the use of ZEBO, but that a combination of pressure and support may. Several authors (in Visscher, 2002) state, for example, that a combination of pressure and support may lead to an increased use of performance feedback. The search for an optimum balance between pressure and support is ongoing. Moreover, the amount of training a school received in the use of ZEBO was not found to influence the use of the ZEBO output. This is probably due to the fact that the majority of respondents answered “I don’t know” to this question.

School organisational characteristics (C) seem to play a crucial role in the use of school self-evaluation results. All four school organisational characteristics influenced the use of ZEBO over the years. Various aspects of the innovation attitude and innovation capacity scale played a role in the use of the ZEBO output in 2003, 2004, and 2006. The most important aspect of the innovation attitude seems to be the degree to which schools staff believes the use of ZEBO will lead to quality improvement. If school staff do not believe that the use of ZEBO will lead to quality improvement, they are less likely to use the results. The most important variable from the innovation capacity scale seems to be “the degree to which the decision to participate in the ZEBO project was taken by the team”. Moreover, extra time and resources should be available for the use of ZEBO. It costs time and money to use the ZEBO output effectively, for example, to discuss the results extensively, to take professional development courses if required, and to implement new teaching methods.

Schools vary greatly in school organisational characteristics. To promote effective school self-evaluation use, some schools need to be restructured in a way which supports and creates interactive processes (Cousins & Leithwood, 1993). This to ensure that self-evaluation results are discussed, understood, a shared meaning is developed, organisational learning occurs and measures are taken to improve the quality of education. If school self-evaluation results are to be used to improve the quality of education, in some schools, school organisational conditions may require change. It may be necessary to, as Weiss (1998a) states, “remove impediments and to supply supportive structures to incorporate and sustain new approaches and actions” (p. 28).

Characteristics of ZEBO (A) as perceived by its users, implementation process features (B) and school organisational characteristics (C) were found to influence the use of ZEBO (D), but part of the variance in ZEBO-use still remains unexplained. This may partly be because of the limited ZEBO use in most schools leading to little variance in the use of ZEBO. In 2003, for example, only 12% of the respondents indicated that ZEBO was instrumentally used. It may also be due to imperfections of this study. Another potential cause for the unexplained variance in ZEBO use is that the framework developed by Visscher (2002) does not include all the variables relevant for studying the use of self-evaluation results. This will be further discussed in the next chapter.

Chapter 7

Conclusion and Discussion

7.1 Background, Research Questions and Method

More and more schools around the world are using some type of school self-evaluation instrument. However, there is a dearth of information about *how* schools use school self-evaluation instruments and what the effects are of the *use* of these self-evaluation instruments. Furthermore, little is known about the critical factors for the successful use of self-evaluation instruments. The main research questions underpinning this study were:

1. How and to what extent do schools use ZEBO?
2. What are the effects of the use of ZEBO?
3. Which factors influence the use of ZEBO?

Self evaluation for this study was defined as *a procedure involving systematic information gathering which is initiated by the school itself and aims to assess the functioning of the school and the attainment of its educational goals for the purposes of supporting decision-making and learning, and for fostering school improvement as a whole.*

ZEBO is a self-evaluation instrument used in Dutch primary schools. In Dutch, the acronym stands for ZelfEvaluatie in het BasisOnderwijs (Self-Evaluation in Primary Schools). It may be used for measuring school and classroom process indicators, which have school effectiveness research as their conceptual background and empirical basis. This instrument comprises four questionnaires: one each for principals, teachers, pupils in grade 3, and for pupils in grades 4-8. Pupils are asked to judge instruction in their class and the topics on which information is collected from pupils are: structured education, adaptive education, classroom climate, and learning time. The topics on which principals are asked to judge school indicators are: co-operation and consultation, pupil care, working environment, educational leadership, professional development of staff and agreement on goals and expectations. Teachers judge instruction in the classroom, as well as the educational organisation at the school level.

An important feature of ZEBO is the comparison of the school and classroom scores for a particular school with the national averages. The ZEBO instrument includes norm-referenced tables of the performance of a representative reference group of Dutch primary schools. Furthermore, school reports compare teachers' scores with those of the principals, and classroom reports compare teachers' scores with those of pupils (Hendriks & Bosker, 2003).

A theoretical framework developed by Visscher (2002) for studying School Performance Feedback Systems (SPFSs) was used as the framework for this study. According to Coe and Visscher (2002a) SPFSs are “information systems external to schools which provide them with confidential information on their performance and functioning as a basis for school self-evaluation” (p. xi). ZEBO is one of those systems which provide schools with confidential information, in the form of school and classroom reports, on their performance and functioning as a basis for school self-evaluation. ZEBO is, however, a school internal system. Based on Visscher’s framework, a new framework was devised for this study.

According to the framework, (see Figure 2.2 in Chapter 2) ZEBO use (D) is influenced by characteristics of ZEBO as perceived by the users of ZEBO (A) (e.g. relevance of output), implementation process features (B) (e.g. clarity of goal), and school organisational characteristics (C) (an example is time and resources for innovation activities). Furthermore ZEBO use (D) may lead to certain effects of ZEBO use (E) on pupil achievement and on several prerequisites for school improvement, for example, didactic methods used by teachers, the functioning of the principal, and the professional development of staff.

Schools administered ZEBO three times, in 2003, 2004, and 2006. To answer the research questions, a questionnaire (Evaluation of ZEBO Questionnaire) was constructed based on a combination of the framework for studying SPFSs developed by Visscher (2002), and the school effectiveness literature. After each of the three administrations of ZEBO, teachers and principals completed the Evaluation of ZEBO Questionnaire. Each time, interviews were then conducted with both teachers and principals to obtain more detailed information on the use of ZEBO, the factors influencing the use of ZEBO, and to validate the results of the Evaluation of ZEBO Questionnaire. In addition, to study the effect of ZEBO use on pupil achievement, two cohorts of pupils were followed using two standardised achievement tests (spelling and mathematics) over the five year period.

In this chapter, a summary of the main results is presented first (section 7.2). Next, some reflections on the research design are offered (section 7.3), followed by the practical and theoretical implications of the results of this study (section 7.4). The chapter concludes with suggestions for further research (section 7.5).

7.2 Conclusions

7.2.1 How and To What Extent Do Schools Use ZEBO?

The findings of this study show that most schools do not make use of the ZEBO output. So far, most of them have been unable to translate the self-evaluation results into measures for improving the quality of education. In 2006, only nine principals and

thirty-one teachers (24%) from ten schools (32%) used the ZEBO output instrumentally. In these ten schools, measures were indeed taken to improve the quality of education based on the ZEBO output. Principals and teachers used the results differently. Principals primarily used the output to make changes at the level of the school organisation. They used the ZEBO output, for example, for writing the school plan, for policy development, and for conducting performance interviews. Teachers used the output to make changes in their teaching.

Although the use of ZEBO was found to have increased in some ways over the years, it still has a long way to go. The results show that although schools went through the procedure of administering ZEBO, the majority of schools did not use the output to improve the quality of education. These findings are in line with the Dutch Inspectorate (2005), who concluded that only one third of primary schools evaluate their functioning on a regular basis and use the findings from a self-evaluation instrument to improve their quality. Since September 1st, 2002, when the new law on the Supervision of Education went into effect, the supervision of schools begins with the results of school self-evaluations, provided the school self-evaluation fulfil the standards set by the Inspectorate. In 2002, quality care in schools was at an early stage of development (Janssens, 2005). Now, five years later, it still seems to be at that stage of development in terms of school self-evaluation processes.

7.2.2 What Are the Intended and Unintended Effects of the Use of ZEBO?

As most schools did not implement ZEBO fully as intended by the instrument's designers, the value of studying the effects of ZEBO use may be questioned. However, although only a small group of schools used the ZEBO output intensively, data from these schools may be compared with the schools which did not make use of the results, in order to investigate whether the use of ZEBO in the schools which did use ZEBO intensively led to an improvement of pupil achievement scores or not. However, no significant relationships were found between the use of ZEBO and pupil achievement. In other words, the use of the ZEBO output has not yet led to greater spelling or mathematics attainment over time. The results also indicate that the use of ZEBO had no negative effects.

ZEBO use was found to have some other effects in the limited number of schools which were using the ZEBO output. The Evaluation of the Use of ZEBO Questionnaire results indicated that ZEBO use led, in these schools, to an improvement in consultation and communication. In addition, quality care became an important subject in team meetings, the mission of the school was discussed more often, and the use of ZEBO led to more discussions on how to improve the school's functioning. Moreover, the didactic methods, the achievement orientation, and professional development improved as a result of the use of ZEBO in 2006. Therefore, although the effects of ZEBO use are limited to date, some effects were found, on important prerequisites for improved pupil achievement in the schools which did use the ZEBO output.

7.2.3 Which Factors Influence the Use of ZEBO?

The finding that most schools did not use the ZEBO output to improve their education provokes another question: Why is this the case? The results of the Evaluation of ZEBO Questionnaires and the interviews suggest that specific characteristics of ZEBO (A), some implementation process features (B), and school organisational characteristics (C) influence the use of the ZEBO (D).

Firstly, three of the ten *characteristics of ZEBO (A)* from the theoretical framework were found to influence ZEBO use (D). Schools, which were more likely to use the ZEBO output to make improvements, were found to be those in which the following characteristics of ZEBO were judged more positively:

- § The degree to which information fits with the needs of the users;
- § The ease of data entry;
- § The time requirements of ZEBO use (these users reported that the use of ZEBO was not time-consuming).

Characteristics of ZEBO were found to influence ZEBO use in the first two evaluations only (in 2003 and 2004). In 2006, these characteristics of ZEBO appear not to have had any further influence on the use of ZEBO. In general, schools which judged the ZEBO characteristics more negatively chose to discontinue using ZEBO. At the end of 2003 and 2004 eleven schools chose to stop using ZEBO and continued with another school self-evaluation instrument.

Contrary to expectations, the other seven characteristics of ZEBO, as perceived by its users, did not influence the use of the ZEBO in 2003, 2004 or 2006. The fact that these seven characteristics of ZEBO, as perceived by its users were, in general, judged positively by the majority of respondents (leading to little variance, and therefore little discrimination between schools which were either high or low users) may explain why these factors were not found to influence ZEBO use.

Secondly, three of the seven *implementation process features (B)* were found to influence ZEBO use (D). The findings suggest that intensive ZEBO users rated the following factors more positively. The degree to which:

- § The goal of using ZEBO is clear to all school staff. In 2003, at the start of this study, the goal of ZEBO was not clear to several teachers, and in 2006, the goal was not clear to several principals. The latter was probably due to the fact that several schools had new principals at that time, who may not have been familiar with ZEBO;
- § School staff feels that they are trained sufficiently in the use of ZEBO. Satisfaction with the training received only played a role in 2004 (when schools used ZEBO for the second time). In 2003 most schools did not use the ZEBO output. In 2004, schools started to use the ZEBO output and they probably realised then that they would benefit from some training in the use of ZEBO. In 2006, teachers and principals indicated that they did not require any further training;

§ The principal encourages the use of the self-evaluation instrument. The principal plays a consistently important role in the use of the self-evaluation results in 2003, and in 2004. In 2006, however, this variable did not explain variance in the use of ZEBO, but closer analysis showed that in that year, 86% of the teachers indicated that the principal continued to encourage the use of ZEBO.

An additional implementation process feature may have also influenced the use of ZEBO. The interview results indicate that it is important that the principal does not only encourage the use of ZEBO but that he or she also takes responsibility for the results: to distribute the results, to explain them where necessary, and to have discussions with school staff on how to use the results to improve the functioning of the school.

Implementation process features, which were found not to influence the use of ZEBO, relate to the pressure to implement ZEBO, the hours of training a school received in the use of ZEBO, and the (amount of and satisfaction with) support. The fact that these factors were not found to influence the use of ZEBO may be due to the lack of variance in these variables. In 2006, for example, only 1.8% of the teachers and principals were not satisfied with the support they received in the use of ZEBO.

Thirdly, all four *school organisational characteristics (C)* were found to have influenced the use of ZEBO. Schools which used the ZEBO output more intensively to make improvements are those in which a positive innovation attitude exists. It was investigated which aspects of the innovation attitude explained variance in the use of ZEBO. Three of the seven items included in the innovation attitude scale were found to influence the use of ZEBO, namely the degree to which:

- § School staff believes that the use of ZEBO will lead to quality improvement;
- § School staff are not beforehand afraid of the changes the use of ZEBO may cause;
- § Teachers feel they are able to influence the measures taken based on ZEBO.

Another school organisational characteristic which consistently influenced the use of ZEBO is the innovation capacity. Schools with a stronger innovation capacity were more likely to have used the ZEBO output. Geijsel (2001) defines the innovation capacity as the capacity of schools to implement innovations in a successful manner. It was investigated which of the items of the innovation capacity scale explained variance in the use of ZEBO. Six of the twelve items included in the innovation capacity scale were found to have influenced the use of ZEBO, including the degree to which:

- § The school monitors the quality of education on a regular basis;
- § School staff experiments with how to improve their education;
- § The principal encourages professional development of teachers;
- § Teachers exchange information on their functioning;
- § The decision to participate in a self-evaluation project was made by the team;

§ The school team is not cohesive. Weaker team cohesion led (only in 2006), in this study, to more ZEBO use in some schools. The reason for this is not clear. It maybe that the ZEBO output was used to (try to) improve team cohesion in these schools. However, this may only happen if an open climate exists.

In addition, schools which had extra time and resources used the ZEBO output more. The interview results show that, for example, in several schools teachers indicated that they required help with the ZEBO administration (specifically with the ZEBO pupil questionnaire). This factor was important in 2004 and remained important in 2006. It costs time and money to use the ZEBO output effectively, for example, to discuss the results extensively, to take professional development courses if required, and to implement new teaching methods.

Another school organisational characteristic which was found to influence ZEBO use is the ZEBO score (number of scales in ZEBO on which the school scored on or above average). If a school scores on or above average on the process indicators measured by ZEBO, there is probably no need, or a reduced need, to use the ZEBO output, since this means that the school scored on or above average on the process indicators measured by ZEBO in comparison with schools from a national sample. To take this into account a variable representing the number of scales on which the school scored on, or above average (the ZEBO score) was constructed. The ZEBO score played a role in 2003 only but not in 2004 and 2006. More schools scored on or above average on more ZEBO scales during the second and third evaluations, making the differences between schools smaller.

The interview results indicate that two additional factors may have influenced the use of ZEBO:

- § The degree to which the self-evaluation results are congruent with the expectations of the principal. If the self-evaluation results are not congruent with the expectations of the principal, the principal may dismiss the results, and this may be the end of the self-evaluation process in the school;
- § The degree to which an open climate exists. School staff must be receptive for critique and issues must be discussable. The two principals interviewed in 2006 hypothesised that this variable could play a role in the use of self-evaluation results. It seems plausible that schools in which school staff is not able to accept criticism, and in which issues are not discussable, the use of self-evaluation results will be difficult, if not impossible.

Part of the variance in ZEBO use remained unexplained. This may partly be because of the limited ZEBO use in most schools leading to little variance in the use of ZEBO. It may also be due to the flaws of this study. Another potential cause for the unexplained variance in ZEBO use is that the framework developed by Visscher (2002) does not include all the variables relevant for studying the use of self-evaluation results. In section 7.4.4, hypotheses around which other variables may influence the use of ZEBO or the use of other SPFSs, are formed.

7.3 Reflections on the Research Design

The study had a longitudinal research design. Two cohorts of pupils were followed over a period of five years. The longitudinal design made it possible to study the effect of the use of ZEB0, over the years, on pupil achievement. A weakness of this study was however that attrition of pupils (e.g. grade retention, transfer to other schools, early transfer to the next grade, and referral to special education) may have biased the results.

The sample for this study was a convenience sample. It was representative for the composition of the pupil population of schools, but not for school denomination nor for school size. The sample was only representative in terms of schools' compositions of pupil populations, which may have resulted in some bias in the findings.

A teacher and principal questionnaire were developed to study the use and effects of ZEB0 as perceived by these two groups of respondents. Due to time restrictions, it was not possible to pilot the developed questionnaire on a large scale. Eight of nine scales were found to be reliable. The questionnaire comprises self-reported data which may have influenced the validity of the results, because teachers and principals may have overrated or underrated their use of the ZEB0 output. In addition, it should be taken into account that, over the years, several schools stopped participating in the project. Furthermore, in 2006, twelve schools chose to postpone administering ZEB0 until 2007. In 2003, 41 principals completed the ZEB0 evaluation questionnaire, and only 25 principals did so in 2006. All in all, this means that the results may present a slightly distorted view.

A final point that must be mentioned here is that this study focussed on the use of a SPFS for school improvement. Schools have much more information available on their quality, which they may also use for school improvement. It is therefore difficult to unambiguously attribute school improvement activities to the introduction of ZEB0.

7.4 Implications of the Study

Despite the reservations described in section 7.3 it is worthwhile to discuss the implications the findings have for schools and for self-evaluation instrument designers, seeking to promote school improvement through self-evaluation. Furthermore, the findings have theoretical implications. The practical and theoretical implications are discussed below.

7.4.1 Implications for Schools

The results of the present study show that the data collected using a school self-evaluation instrument, such as ZEB0, should fit with the needs of the users, the data entering options should be user-friendly, and working with the instrument should not take too much time. This implies that if schools are considering using a new self-evaluation instrument, it is important that school staff select an instrument which fits with their preferences.

School staff should be involved in the entire school self-evaluation process, from choosing a self-evaluation instrument, to using the self-evaluation results to make improvements. Teachers, for example, should feel that they are able to influence measures taken based on the results of the self-evaluation. This encourages ownership of the (results of the) self-evaluation, which is a factor known to play a role in the use of performance feedback (Davies & Rudd, 2001; Tymms & Albone, 2002; Row et al., 2002).

The findings further show that it is important to establish clarity about the purpose of the self-evaluation, about how to conduct the self-evaluation, and how the results will be used (for example, for accountability purposes or for school improvement). The principal may play an important role in this light. Sometimes principals must deal with a negative attitude from teachers towards school self-evaluation. Schools, in which the principal actively and enthusiastically encourages and supports the self-evaluation, are more likely to use the evaluation results. Some schools lack an active, encouraging and supporting principal. In these schools self-evaluation is probably more difficult, but not impossible. Somebody must take responsibility for the self-evaluation process. This person must not necessarily be the principal. Devos and Verhoeven (2003) recommend, in this light, the involvement of powerful external stakeholders such as local school authorities, which may provide an important counter-weight to the internal self-evaluation process.

Furthermore, in this study more innovative schools were more likely to use self-evaluation results to improve their functioning, as they have principals who will encourage the use of self-evaluation results, and staff who participate in decision making. Aspects of the innovation capacity found to influence the use of self-evaluation results are mainly teacher related variables, such as the degree to which teachers exchange information on their functioning, and the extent to which they feel that the principal encourages their professional development. This implies (like Geijsel (2001) also suggests) that to realise change in the classroom (e.g. to make sure that teachers are using the self-evaluation results to improve their functioning) one must focus especially on the teacher, and less on the entire school as a unit of change.

The findings of this study further suggest that schools which have more resources (e.g. time, money, manpower) and training at their disposal for implementing and using (the results of) a self-evaluation instrument, are more likely to use the self-evaluation results to improve their functioning. Teachers and principals face considerable challenges, both with having to interpret the self-evaluation results, and with developing the skills and competencies required to use to self-evaluation results to improve their functioning. Training and educating principals and teachers in this area could compensate for their lack of skills, knowledge and competencies.

A question which arises here is who should provide schools with these resources and training? One could argue that since the school itself is responsible for its quality, the school itself must acquire the required resources and training. Scheerens (2006), for example, found that training and support in the form of professional development programs and participating in a network of schools, led to higher levels of self-evaluation use. Instead of focussing on on-the-job training, one may also want to consider making evaluation skill training part of the curriculum of teacher training colleges.

7.4.2 Implications for School Self-Evaluation Designers

Although the results show that the use of self-evaluation results is slowly increasing, using a self-evaluation instrument effectively proves to be difficult for most schools. The difficulty appears to be associated with the steps following the collection of indicators on school quality rather than the actual administration of a self-evaluation instrument. The ZEBO instrument does not provide users with guidelines for potential improvement actions if a school scores below average. A method of promoting the use of self-evaluation results may therefore be the inclusion of guidelines on how to use the self-evaluation results, as well as examples of actions a school may undertake in order to make improvements.

In addition, the fact that several characteristics of self-evaluation instruments, such as the effort it takes to enter data and to work with the instrument, influence the use of the self-evaluation results implies that designers face a major challenge in developing user-friendly self-evaluation instruments.

Another challenge for school self-evaluation developers lies in ensuring that the instrument provides the user with rapid and frequent feedback. Several authors found that rapid, frequent, and individualised feedback is associated with stronger feedback effects (Yeh, 2006; Kluger & DeNisi, 1996; Fitz-Gibbons, 1996, in Coe & Visscher, 2002a).

In addition to implications for schools and designers, the results of the present study also have implications for theories on performance feedback, which are presented in the next section.

7.4.3 Theoretical Implications

A theoretical framework for studying SPFSs developed by Visscher (2002) was used as the basis for this study of the use of ZEBO. Although in this study, the use of only one specific internal SPFS was investigated, some theoretical implications may be derived from the results. The SPFS framework has proven to be very usable for studying the use of an instrument such as ZEBO. However, not all variance in the use of the performance feedback could be explained. This may have had various causes, for example, the instruments used for data collection. However, it may also imply that the framework does not cover all the factors which play a role in the use of performance feedback. Below, some more general theoretical implications of the results of this study are discussed.

Firstly, the results show that *the use of performance feedback* was limited. The performance feedback was used to discuss the quality of the school, however, in most schools no measures were taken to improve the quality of the school. In other studies, the use of performance feedback obtained using school self-evaluation instruments was also found to be limited (Janssens, 2005; De Inspectie van het Onderwijs, 2005; Hofman, Dijkstra, Hofman, & De Boom, 2004; Blok, Slegers, & Karsten, 2005).

A factor not included in the use of the SPFS in this framework, is the comparison of the SPFS results and the educational goals of the schools. Ofsted (2006) describes asking how well a school has performed in relation to its targets, as a crucial first stage in the school self-evaluation process. Schools must determine to what extent their educational goals have been realised. The use of self-evaluation results also includes comparing the results to the goals set, if necessary, taking actions to reduce discrepancies between the results and the goals, and setting new goals and objectives for school improvement based on the performance feedback.

Secondly, regarding the *effects of performance feedback*, Kluger and DeNisi (1996) conclude that the effects of feedback are complex and sometimes even harmful. The results of this study suggest that the effects of feedback are indeed complex, but not necessarily harmful. No negative or harmful effects of the ZEBO output were reported. Moreover, feedback had an effect on several important prerequisites for improved pupil achievement in some of the schools in this study.

Thirdly, the results of this study suggest that characteristics of the SPFS, as perceived by its users, implementation process features, and school organisational characteristics may influence the use of the performance feedback (see section 7.2.3). Regarding the *characteristics of a SPFS as perceived by its users*, the degree to which the feedback fits with the needs of the users was found to influence the use of the performance feedback twice in this study. It seems likely that this variable will also influence the use of performance feedback generated by other SPFSs.

The characteristics which influence the use of SPFS results probably differ for various SPFSs. Therefore, searching for common characteristics of SPFSs which promote the use of performance feedback independently from the contexts in which they are used is recommended. One may think of variables such as the degree to which the instrument is user-friendly.

Furthermore, in line with Kluger and DeNisi (1996), cues which draw attention to the self (the personality of a person) may cause the feedback to be disregarded. ZEBO provides users with feedback on their current functioning in comparison with a national average, which may direct too much attention to the self. In line with Coe (1998a, in Coe, 2002), it is therefore recommended that performance feedback should focus more on the task and on people's current performance, relative to their past performance, instead of directing attention to comparisons of their performance with that of others.

In addition, feedback should focus on the task and include cues on how to improve upon that task (e.g. on how to improve one's performance, or on how to decrease the discrepancy between current performance and the standard) (Kluger & DeNisi, 1996). The lack of specific cues in the ZEBO output may have contributed to the limited use of the feedback.

Implementation process features included in the framework, such as training, the clarity of the goal of the SPFS, and the encouragement of the use of the SPFS, were found to influence the use of the performance feedback generated by ZEBO. These variables may also influence the use of other SPFSs. The results of this study, for example, indicate that performance feedback is more efficient, if the goal of the feedback is clear to the recipient, prior to receiving the feedback.

School organisational characteristics in the framework, including the score on the SPFS, time and resources available for the use of the SPFS, the innovation attitude of schools staff, and the innovation capacity, were all found to influence the use of performance feedback in this study, and may also influence the use of other SPFSs.

In addition, the extent to which schools have set clear educational goals and a clear mission may also influence the use of SPFSs. Blok et al (2005) state that goal setting is an important step in the quality care process. Schools which have set clear goals may have fewer difficulties in using self-evaluation results. These schools may compare the self-evaluation results to the goals set beforehand and take measures to decrease discrepancies if necessary.

Another school organisational characteristic, as of yet not included in the SPFS framework, which may influence the use of performance feedback, is the degree to which other sources of information are seriously competing for the attention of the users (Leithwood et al., 2001). Schools receive a lot of information on their functioning, such as self-evaluation results, but also school inspection reports and assessment data. Schools must decide what information they will use in their quality care process and may, for example, prioritise the use of inspection information. A possible way of making school self-evaluation results more competitive may be by involving all school staff in the planning, and carrying out of the self-evaluation, to increase commitment to the evaluation process and the evaluation results, as recommended by Cousins and Leithwood (1986), which may in turn lead to school staff prioritising self-evaluation results over other data.

Kluger and DeNisi (1996) conclude in their review of the impact of feedback interventions on performance that feedback effects also depend on *personality variables*. Based on several studies (Kluger & DeNisi, 1996; Scheerens, 1996; Tokar, Fischer, and Subich, 1998; Geijsel, 2001; Hofman et al., 2004) it is expected that characteristics of the individual self-evaluation users, such as motivation, the locus of control, and feelings of uncertainty, influence the use of self-evaluation results. Hofman

et al. (2004) found, for example, that the motivation of teachers plays an important role in quality care processes within schools. Tokar et al. (1998) found that people with a higher internal locus of control (attributing success or failure to themselves) fare better in change related processes. This may imply that schools with a majority of school staff with a high internal locus of control do better in educational change and thus are more inclined to use the SPFS results, to make changes to improve the quality of education than schools with a low(er) internal locus of control. Geijsel (2001) found in her research into conditions fostering the implementation of innovations, that feelings of uncertainty in teachers negatively influenced the implementation of educational innovations. SPFS use is another example of a new challenge teachers must deal with, in an ongoing stream of new situations and challenges. Feelings of uncertainty may hinder the use of SPFS results.

The results of this study provide some insights into the use of performance feedback, its effects, and some insights into what circumstances promote effective performance feedback. However, the use and effects of performance feedback, as well as the moderators of performance feedback, are still far from fully understood. More research is therefore required. In section 7.5 some suggestions for further research are presented.

7.5 Suggestions for Further Research

Based on the results of this study several suggestions for further research may be made. Firstly, the questions “How do schools use self-evaluation results?”, “What are the effects of using a school self-evaluation instrument?”, and “Which factors influence the use of self-evaluation results?”, remain important questions for research, especially since the results of this study show that schools are just starting to use self-evaluation results. Although this study covered a period of five years, this was still not enough for most schools to use school self-evaluation results effectively. It would be interesting to further follow these schools. Will the use of school self-evaluation results grow further, or will it stagnate or even decrease? Will the same variables continue to influence the use of self-evaluation results when schools start to make more and more use of the results? Moreover, this study only considered the use of one specific school self-evaluation instrument. Research is required into the use of other school self-evaluation instruments, and in other contexts.

Another interesting research question would be “Are the factors which are decisive for successful school self-evaluation use, the same for different school self-evaluation instruments and for different countries?” The results of a study into the use of performance feedback conducted by Schildkamp and Teddlie (submitted for publication) seem to indicate that several factors which are important in the Netherlands are also important in the United States. It would be interesting to study further the use of performance feedback in more countries, using a research design similar to the one used in the present study. Since it appears that school self-evaluation is here to stay, and that schools and policy

makers all over the world are putting time, money and effort into school self-evaluation systems and SPFSs, it is suggested that it would be worthwhile to find and encourage the factors which influence the use of performance feedback generated by these systems. If, for example, performance feedback is to be used to improve the quality of education, in some schools, school organisational conditions must be changed.

An alternative for studying the use of SPFSs or school self-evaluation instruments would be a case study research design. Such a design would be more appropriate to answer in-depth questions, such as “How do schools make sense of the data”, and “How do school staff decide which measures to take, based on the data?” Yin (1994) states that case studies are the preferred strategy when “how questions” are being posed, when the researcher has little control over the events, and when the focus is on a contemporary phenomenon within a real-life context. Case studies may help to clarify how schools use performance feedback over a longer period of time, by not only conducting interviews and administering questionnaires, but also by observing the whole self-evaluation process in individual schools, including, for example, the discussion of the performance feedback.

Furthermore, an aspect of the school self-evaluation process, which deserves specific attention, is training and support. The results of this study show that although school self-evaluation may have several benefits for schools, it is a complex process, which most schools are not capable of handling unsupported. Macbeath, Schratz, Meuret and Jakobsen (2002) describe the role of training and support provided by a critical friend in school self-evaluation. A critical friend is a person who may bring an outside perspective, a reference point, and a connection with a wider field of knowledge into the school. The critical friend should have a critical attitude to the schools so as to challenge its practice, but should also give school staff unconditional support so that they feel accepted and listened to. Macbeath et al. (2002) do not clarify who should “play” the role of the critical friend. Possibly a person from the school advisory service may act as a critical friend. Hofman et al. (2004) found, for example, that schools valued training and support provided by the school advisory service in their quality care process. Furthermore, training and support in the form of professional development programs and participating in a network of schools may lead to higher levels of self-evaluation use, as was found by Scheerens (2006). It would be interesting to study the influence of different kinds of training and support on the use of school self-evaluation instruments. “What kind of training and support do schools require in the use of self-evaluation results?”, and “Who should provide the training and support?” are important questions.

Next to school self-evaluation results and school inspection results, schools have yet other data available to assess their functioning and improve their performance, such as pupil assessment data. Another interesting question would be whether schools use these data to improve their functioning, and if yes, precisely what kinds of data schools use to assess and improve their functioning. Ingram, Seashore, Louis and Schroeder (2004)

found, for example, that not all decisions made in schools are data-based decisions. Their analyses showed that approximately 40% of the remarks included a description of using systematic data for decision making, an equivalent proportion of the remarks reflected the use of anecdotal information, experience, or intuition to make decisions, and about 15% of the remarks described using a combination of some type of systematic data and some type of non-systematic data such as anecdotes. Much may thus be gained by helping schools to make more data-based decisions. Wayman and Stringfield (2006) also pose interesting questions, in this regard, “What forms of data use are most beneficial for educators at various levels of the school system?”, and “What practices and support may be provided to, and by, teachers and principals in improving effective and enhanced data use?”

In conclusion, this study provides more insight into the use and effects of school self-evaluation systems, and into the factors influencing the use of a school self-evaluation instrument. However, it also raises new interesting questions. Since quality care and the development of school evaluation instruments are receiving much attention around the globe, the justifications for using self-evaluation instruments are plausible. Thousands of schools around the world have voluntarily implemented them (an example includes the SPFSs from institutions such as the “Curriculum Evaluation and Management Centre” in the United Kingdom, used in many countries) (Schildkamp & Visscher, submitted for publication). More research into the use of these systems, in the ways previously proposed, is recommended in order to gain better insight into how schools may benefit more fully from the internal or external SPFS they already use.

References

- Baarda, D. B., Goede, M. P. M. de., & Teunissen, J. (1997). *Basisboek kwalitatief onderzoek [Handbook for qualitative research]*. Houten: Educatieve Partners Nederland BV.
- Berg, van den R., & Vandenberghe, R. (1999). *Succesvol leiding geven aan onderwijsinnovaties. Investeren in mensen [Successfully leading educational innovations: Investing in people]*. Alphen aan den Rijn: Samsom.
- Berg, van den R., Vandenberghe, R., & Slegers, P. (1999). Management of innovations from a cultural-individual perspective. *School Effectiveness and School Improvement*, (10)3, 321-351.
- Biesheuvel, M., & Flim, G. (2001). Het vermogen van leerkrachten de intelligentie van kleuters in te schatten [The ability of teachers to assess pupils' intelligence]. In W. J. Van der Linden, & S. Oosterloo (Eds.), *Proceedings van het symposium in het kader van het vak onderzoeksopdracht (196037)*. Enschede: Faculty of educational science and technology. Department of educational measurement and data-analysis. University of Twente.
- Blok, H., Slegers, P., & Karsten, S. (2005). *Schoolzelfevaluatie in het basisonderwijs: Een terugblik op de zelfevaluatiefase van Ziezo [School self-evaluation in primary education: Looking back at the phase of selfevaluation of ZIEZO]*. Amsterdam: SCO-Kohnstamm Instituut.
- Bosker, R. J. (2001). Prikkels tot kwaliteitsverbetering [Stimulus for quality improvement]. In R. J. Bosker (Ed.), *Kwaliteitszorg. Onderwijskundig lexicon III* (pp. 13-26). Alphen aan den Rijn: Samsom.
- Bosker, R. J., & Hendriks, M. A. (1997). *Betrouwbaarheid, validiteit en bruikbaarheid van een instrumentarium ten behoeve van schoolzelfevaluatie in het basisonderwijs [Reliability, validity, and usability of a school self-evaluation instrument for primary education]*. Enschede: OCTO.
- Bruggen, J. van, & Mertens, F. (2001). Inspecteren voor openbaarheid vanuit bredere kwaliteitsnoties. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: Scholen op rapport [The eye of the nation: School reports]* (pp. 208-218). Assen: Koninklijke Van Gorcum.
- Burke Johnson, R., & Onwuegbuzie, A. J. (2004). Mixed methods research paradigm: A research whose time has come, *Educational Researcher*, 33(7), 14-26.
- 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: Swets & Zeitlinger B.V.
- Coe, R., & Visscher, A. J. (2002a). Introduction. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (pp. xi-xix). Lisse: Swets & Zeitlinger B.V.
- Coe, R., & Visscher, A. J., (2002b). Drawing up the balance sheet for school performance feedback systems. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (221-254). Lisse: Swets & Zeitlinger B.V.
- Cousins, B., & Leithwood, K. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331-364.
- Cousins, B., & Leithwood, K. (1993). Enhancing knowledge utilization as a strategy for school improvement. *Knowledge: Creation, diffusion, utilization*, 14(3), 305-333.
- Cremers-van Wees, L. M. C. M., Rekveld, I. J., Brandsma, H. P., & Bosker, R. J. (1996a). *Instrumenten voor zelfevaluatie: Inventarisatie en beschrijving [Instruments for school self-evaluation: Inventory and description]*. Enschede: Universiteit Twente, Onderzoek Centrum Toegepaste Onderwijskunde.
- Cremers-van Wees, L. M. C. M., Rekveld, I. J., Brandsma, H. P., & Bosker, R. J. (1996b). *Instrumenten voor zelfevaluatie: Beschrijving van 31 instrumenten [Instruments for school self-evaluation: Description of 31 instruments]*. Enschede: Universiteit Twente, Onderzoek Centrum Toegepaste Onderwijskunde.
- Davies, D., & Rudd, P. (2001). *Evaluating school self-evaluation*. Berkshire: National Foundation for Educational Research.
- Deckers, J., & Jacobs, J. (1994). *Zelfevaluatie en visitatie. Instrumenten voor kwaliteitszorg in de autonome school [Self-evaluation and external review: Instruments for quality care in the autonomous school]*. Houten: Educatieve partners Nederland bv.
- Devos, G., & Verhoeven, J. C. (2003). School self-evaluation – Conditions and caveats. The case of secondary schools. *Educational Management & Administration*, 31(4), 403-420.
- Doolaard, S., & Karstanje, P. (2001). Gebruik van publieke prestatie-indicatoren voor schoolverbetering [The use of public performance indicators for school improvement]. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie. Standaarden voor de publicatie van schoolprestaties [The eye of the nation: Standards for the publication of school performance]* (pp. 155-175). Assen: Koninklijke Van Gorcum BV.
- Eggens, T. J. H. M., & Sanders, P. F. (Eds.). *Psychometrie in de praktijk [Psychometrics in practice]*. Arnhem: CITO.

- Earl, L., & Fullan, M (2003). Using data in leadership for learning. *Cambridge Journal of Education*, 33(3), 383-394.
- Evers, A., Vliet-Mulder, J. C. van, Resing, W. C. M., Starren, J. C. M. G., Alphen De Veer, R. J. van, & Boxtel, H. van. *COTAN testboek voor het onderwijs [COTAN test book for education]*. Amsterdam: NDC-Boom.
- Field, A. (2003). *Discovering statistics using SPSS for windows*. London: SAGE Publications Ltd.
- Follman, J. (1991). Teachers' estimates of pupils' IQS and pupils' tested IQS. *Psychological reports*, 69, 350-350.
- Fullan, M. G (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Geijsel, F. (2001). *Schools and innovations. Conditions fostering the implementation of educational innovations*. Nijmegen: Nijmegen University Press.
- Geijsel, F. P., & Krüger, M. L. (2005). Leren van onderzoek: De benutting van datafeedback ten behoeve van schoolontwikkeling [Learning from research: The use of performance feedback for school development]. *Pedagogische studiën*, 82, 327-342.
- Geijsel, 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.
- Geijsel, F., Van Den Berg, R., & Slegers, P. (1999). The innovative capacity of schools in primary education: a qualitative study. *Qualitative studies in education*, volume 12, 2, 175-191.
- Gennip, van J. (1991). *Veranderingscapaciteiten van basisscholen: Onderzoek naar de functie van veranderingscapaciteiten binnen het innovatieproces basisonderwijs [The innovation capacity of primary schools: Research into the function of innovation capacity within primary education]*. Nijmegen: Instituut voor Toegepaste Sociale Wetenschappen.
- Gennip, van J., & Claessen, van H. (1993). Veranderen of niet veranderen: Over het imago van scholen [Change or not to change: About the image of schools]. *School en Begeleiding*, 9(37), 35-38.
- Ghere, G., King, J. A., Stevahn, L., & Minnema, J. (2006). A professional development unit for reflecting on program evaluator competencies. *American Journal of Evaluation*, 27(1), 108-123.
- Grasso, P. G. (2003). What makes an evaluation useful? Reflections from experience in large organizations. *American Journal of Evaluation*, 24(4), 507-514.

- 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* (143-162). Lisse: Swets & Zeitlinger B.V.
- Griffith, J. (2002). Is quality/effectiveness an empirically demonstrable school attribute? Statistical aids for determining appropriate level of analysis. *School Effectiveness and School Improvement*, 13(1), 91-122.
- Grift, W. van de (2001). Evaluatie van de kwaliteit van het basisonderwijs door de inspectie. In A. B. Dijkstra, S. Karsten, R. Veenstra, & A. J. Visscher (Eds.), *Het oog der natie: Scholen op rapport [The eye of the nation: School reports]* (pp. 219-235). Assen: Koninklijke Van Gorcum.
- 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: Samsom.
- 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: Swets & Zeitlinger B.V.
- Hendriks, M., & Bosker, R. (2003). *ZEBO instrument voor zelfevaluatie in het basisonderwijs. Handleiding bij een geautomatiseerd hulpmiddel voor kwaliteitszorg in basisscholen*. Enschede: Twente University Press.
- Heus, P. de., Leeden, R. van der., & 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: Elsevier Bedrijfsinformatie bv.
- Hoeben, W. Th. J. G. (1995). Integrale kwaliteitszorg [Integral quality care]. In J. Lowyck, & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals [Educational science: A knowledge basis for professionals]* (pp. 287-314). Groningen: Wolters-Noordhoff.
- Hofman, R. H., Dijkstra, N. J., Hofman, & De Boom, J. (2004). *Kwaliteitszorg in het primair onderwijs. BOPO-Project (412-02-010). Deelstudie 1: Peiling 2003/2004 [Quality care in primary education. BOPO-project (412-02-010. Study 1, over the years 2003/2004]*. Groningen: Gion.
- Houtveen, A. A. M., Booij, N., Jong, R. de, & Grift, W. J. C. M. van de (1996). Adaptief onderwijs en leerlingresultaten [Adaptive education and student achievement]. *Pedagogische Studiën*, 73, 422-433.
- Houtveen, A. A. M., Booij, N., Jong, R. de, & Grift, W.J.C.M. van de (1999). Adaptive instruction and pupil achievement. *School effectiveness and school improvement*, 10(2), 172-192.

- Hox, J. (2002). *Multilevel analyses. Techniques and application*. New Jersey: Lawrence Erlbaum Associates.
- Ingram, D., Seashore Louis, K., & Schroeder, R. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106(6), 1258-1287.
- Inspectie van het Onderwijs (2002). *Onderwijsverslag over het jaar 2001 [Education Report for 2001]*. Den Haag: SDU.
- Inspectie van het onderwijs (2005). *Toezichtkader primair onderwijs 2005. Inhoud en werkwijze van het inspectietoezicht [Inspection framework primary education 2005. Content and method of inspection]*. Utrecht: Inspectie van het onderwijs.
- Janssens, F. J. G. (2005). *Toezicht in discussie. Over onderwijstoezicht en educational governance [Discussing inspection: About inspection and educational governance]*. Enschede: Universiteit Twente.
- 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.
- Kluger, A. N., & 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(2), 254-284.
- Kyriakides, L., & Campbell, R. J. (2004). School self-evaluation and school improvement: A critique of values and procedures. *Studies in Educational Evaluation*, 30, pp. 23-36.
- Ledoux, G., Overmaat, M., & Koopman, P. (1997). *Kwaliteitszorg in het primair onderwijs. Secundaire analyses op de PRIMA-cohort bestanden [Quality care in primary education: Secondary analyses on the PRIMIS-cohort data]*. Amsterdam: SCO-Kohnstamm Instituut.
- Leithwood, K., Jantzi, D., & Steinbach, R. (1999). *Changing leadership for changing times*. Philadelphia: Open University Press.
- Leithwood, K., Aitken, R., & Jantzi, D. (2001). *Making schools smarter. A system for monitoring school and district progress*. Thousand Oaks: Corwin Press, Inc.
- Litosseliti, L. (2003). *Using focus groups in research*. London: Continuum.
- MacBeath, J., & Mortimore, P. (2001). School effectiveness and school improvement: The story so far. In J. MacBeath, & P. Mortimore. *Improving school effectiveness* (pp 1-21). Buckingham and Philadelphia: Open University Press.
- MacBeath, J., Schratz, M., Meuret, D., & Jakobsen, L. (2000). *Self-evaluation in European schools*. London: Routledge Falmer.

- MacCallum, R., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490-504.
- Malin, A., & Linnakylä, P. (2001). Multilevel modelling in repeated measures of the quality of Finnish school life. *Scandinavian Journal of Educational Research*, *45*(2), 145-166.
- Marx, E. C. H., De Vries, O. A., Veenman, S. A. M., & Slegers, P. (1995). School en klas [Teacher and class]. In J. Lowyck, & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals [Educational science: A knowledge base for professionals]* (pp. 43-77). Groningen: Wolters-Noordhoff.
- Meijer, P. C., Verloop, N., Beijaard, D. (2002). Multi-method triangulation in a qualitative study on teachers' practical knowledge: An attempt to increase internal validity. *Quality & Quantity*, *36*(2), 145-167.
- Mertens, D. (1998). *Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches*. Thousand Oaks, CA: Sage.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks: Sage.
- Ministerie van Onderwijs, Cultuur & Wetenschappen (1999). *Variëteit en waarborg: Voorstellen voor de ontwikkeling van het toezicht op onderwijs [Diversity and a guarantee: Proposals for the development of the supervision of education]*. Zoetermeer: Ministerie van Onderwijs, Cultuur & Wetenschappen.
- Ministerie van Onderwijs, Cultuur & Wetenschappen; Ministerie van Landbouw, Natuur en Visserij (2000-2002). *Wet op het onderwijstoezicht [Education supervision act]*. Den Haag: SDU.
- Moelands, H., Engelen, R., Ouborg, M., Hendriks, M., Doolaard, S., Fond Lam, J., Bosker, R., Frowijn, R., Goffree, F., & Scheerens, J. (2000). *Zelfevaluatie basisonderwijs (ZEBO). Naar een geïntegreerd instrumentarium. Eindrapport van het gezamenlijke project van Cito, OCTO en SLO [Self-evaluation in primary education. Final report of the collaborative project of OCTO, SLO and OCTO]*. Universiteit Twente: OCTO.
- Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology*, *22*, pp. 129-152.
- Mortimore, P. (1998). *The road to improvement. Reflections on school effectiveness*. Lisse: Swets & Zeitlinger B.V.
- Ofsted (2006). *Improving performance through school self-evaluation and improvement planning*. Manchester: Ofsted.
- Olthof, A., Emmerik, R., & Troost, F. (1993). *Integrale kwaliteitszorg in het voortgezet onderwijs. Parels in stromend water [Quality care in secondary education]*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.

- Petegem, P. Van (1998a). *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: Uitgeverij Acco.
- Petegem, P. Van (1998b). Zelfevaluatie als methodiek voor beleidsvoering scholen [Self-evaluation for school policy development]. *MESO Magazine*, 103, 17-25.
- Petegem, P. Van (2001). Kwaliteitszorg of (zelf-)evaluatie? [Quality care or (self)-evaluation?]. In R. J. Bosker (Ed.), *Kwaliteitszorg. Onderwijskundig lexicon III* (pp. 27-46). Alphen aan den Rijn: Samsom.
- Petegem, P. Van, & Vanhoof, J. (2002a). Feedback of indicators to schools. In J. Scheerens, & M. Hendriks (Eds.), *Benchmarking the quality of education* (131-162). Enschede: Universiteit Twente.
- Petegem, P. Van, & Vanhoof, J. (2002b). Flemish school principals on their use of individual school feedback. In J. Scheerens, & M. Hendriks (Eds.), *Benchmarking the quality of education*. (263-273). Enschede: Universiteit Twente.
- Petegem, P. Van., & Vanhoof, J. (2004). Feedback over schoolprestatie-indicatoren als strategisch instrument voor schoolontwikkeling? Lessen uit twee Vlaamse cases [Feedback on performance indicators as a strategic instrument for school development? Lessons from two Flemish cases]. *Pedagogische Studiën*, 5, 338-353.
- Petegem, P. Van., Vanhoof, J., Daeam, F., & Mahieu, P (2005). Publishing information on individual schools. *Educational Research and Evaluation*, 11(1), 45-60.
- Porter, A. C. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, 13(1), 13-29.
- Rasbash, J., Steele, F., & Browne, W. (2004). *A user's guide to MLwiN*. London: Centre for multilevel modelling.
- Renkema, G. (2002). *Kwaliteitszorg, wat levert het op? [Quality assurance – What are the gains?]*. Basisschoolmanagement, 15(8). Alphen aan den Rhijn: Samsom.
- Reynolds, D., Hopkins, D., Potter, D., & Chapman, C. (2002). *School improvement for schools facing challenging circumstances: A review of research and practice*. London: OFSTED. Retrieved from www.kwaliteitsring.nl .
- Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation. A systematic approach*. Thousands Oaks: SAGE Publications.
- 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* (163-190). Lisse: Swets & Zeitlinger B.V.
- Saunders, L (1999). Who or what is 'self'-evaluation for? *School Effectiveness and School Improvement*, 10(4), 414-429.

- 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(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: Cassel.
- Scheerens, J. (1996). *Beoordeling en evaluatie in het kader van kwaliteitszorg in het onderwijs [Judgement and evaluation for quality care in education]*. In J. Scheerens (Ed.), *Onderwijskundig lexicon. Kwaliteitszorg in het onderwijs*. Alphen aan den Rijn: Samsom H.D. Tjeenk Willink.
- Scheerens, J. (1997). *De bevordering van schooleffectiviteit in het basisonderwijs: Mogelijkheden tot "flankerend beleid" bij klassenverkleining [The enhancement of school effectiveness in primary education: Opportunities for policy in reducing class size]*. Enschede: Faculteit der Toegepaste Onderwijskunde, Universiteit Twente.
- Scheerens, J. (1999). Concepts and theories of school effectiveness. In A. J. Visscher (Ed.), *Managing schools towards high performance. Linking school management theory to the school effectiveness knowledge base*. Lisse: Swets & Zeitlinger B.V.
- Scheerens, J. (2006). *Creating a culture for school self-evaluation; Theoretical and practical considerations*. Paper presented at the Conference on "The prospects and possibilities for creating learning societies in an age of uncertainty and insecurity". Athens, September 29 – October 1.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Elsevier Science Ltd.
- Scheerens, J., Glas, C., & Thomas, S. M (2003). *Educational evaluation, assesement, and monitoring*, Lisse: Swets & Zeitlinger.
- Schildkamp, K., & Teddlie, C. (submitted for publication). *School performance feedback systems in the USA and in the Netherlands: A comparison*. Paper submitted for publication.
- Schildkamp, K., & Visscher, A. J. (submitted for publication). *Factors influencing the utilization of a school self-evaluation instrument*. Paper submitted for publication.
- Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 8(2&3), 277-310.
- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage.

- Steckler, A., McLeroy, K. R., Goodman, R. M., Bird, S. T., & McCormick, L. (1992). Toward integrating qualitative and quantitative methods: An introduction. *Health Education Quarterly*, 19(1), 1-8.
- Strickley, A. (2004). Factors affecting the use of MIS as a tool for informing and evaluating teaching and learning in schools. *Education and Information Technologies*, 9(1), 47-66.
- Teddle, 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: Swets & Zeitlinger B.V.
- Tellegen, P., Winkel, M., Wijnberg-Williams, B., & Laros, J. (1998). *Snijders-Oomen niet-verbale intelligentietest: SON-R 2½-7. Handleiding en verantwoording [Snijders-Oomen non-verbal intelligence test: SON-R2½-7. Manual]*. Lisse: Swets & Zeitlinger.
- The Standing International Conference of Central and General Inspectorates of Education (2003). *Effective school self-evaluation*. County Down: SICI
- Tockar, D. M., Fischer, A. R., & Mezydlo Subich, L. (1998). Personality and vocational behavior. A selective review of the literature, 1993-1997. *Journal of Vocational Behavior*, 53, 115-153.
- Tymms, P., & Albone, S. (2002). Performance indicators in primary schools. In A. J. Visscher, & R. Coe (Eds.), *School improvement through performance feedback* (191-218). Lisse: Swets & Zeitlinger B.V.
- Verloop, N. (1995). De leraar [The teacher]. In J. Lowyck, & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals [Educational science: A knowledge base for professionals]* (pp. 109-150). Groningen: Wolters-Noordhoff.
- Verloop, N., & Schoot, F. van der (1995). Didactische evaluatie [Didactic evaluation]. In J. Lowyck, & N. Verloop (Eds.), *Onderwijskunde: Een kennisbasis voor professionals [Educational science: A knowledge base for professionals]* (pp. 249-284). Groningen: Wolters-Noordhoff.
- 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: Swets & Zeitlinger B.V.
- Visscher, A. J., & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement*, 14(3), 321-350.
- Vlug, K. F. M. (1997). Because every pupil counts: The success of the pupil monitoring system in the Netherlands. *Education and Information Technologies*, 2, 287-306.
- Wayman, J. C., & Stringfield, S. (2006). Data use for school improvement: School practices and research perspectives. *American Journal of Education*, 112, 463-468.

- Webb, R., Vulliamy, G., Häkkinen, K., & Hämäläinen, S. (1998). External inspections or school self-evaluation? A comparative analysis of policy and practice in primary schools in England. *British Educational Research Journal*, 24(5), 539-566.
- 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*, 7, 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.
- Wild, K-P. (1993). Teacher-judgments and the identification of gifted pupils. *Studies in Educational Psychology*, 5, 1-22.
- 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.
- Yeh, S. S. (2006). Can rapid assessment moderate the consequences of high-stakes testing? *Education and Urban Society*, 39(1), 91-112.
- Yin, R. K. (1994). *Case study research. Design and methods*. London: Thousand Oaks.

English Summary

Introduction and Research Questions

More and more schools around the world are using some type of school self-evaluation instrument. However, there is a dearth of information about *how* schools use school self-evaluation instruments and what the effects of the use of these instruments are. Furthermore, little is known about the critical factors for the successful use of self-evaluation instruments. The main research questions underpinning this study were:

1. How and to what extent do schools use ZEBO?
2. What are the effects of the use of ZEBO?
3. Which factors influence the use of ZEBO?

Self evaluation for this study was defined as *a procedure involving systematic information gathering which is initiated by the school itself and aims to assess the functioning of the school and the attainment of its educational goals for the purposes of supporting decision-making and learning and for fostering school improvement as a whole.*

ZEBO is a self-evaluation instrument used in Dutch primary schools. In Dutch, the acronym stands for *ZelfEvaluatie in het BasisOnderwijs* (Self-Evaluation in Primary Schools). It can be used for measuring school and classroom process indicators, which have school effectiveness research as their conceptual background and empirical basis. This instrument comprises four questionnaires: one for principals, one for teachers, one for pupils in grade 3, and one for pupils in grades 4-8. Pupils are asked to judge instruction in their class and the topics on which information is collected from pupils are: structured education, adaptive education, classroom climate, and learning time. The topics on which principals are asked to judge school indicators are: co-operation and consultation, pupil care, working environment, educational leadership, professional development of staff and agreement on goals and expectations. Teachers judge instruction in the classroom, as well as the educational organization at the school level.

An important feature of ZEBO is the comparison of the school and classroom scores for a particular school with the national averages. The ZEBO instrument includes norm-referenced tables of the performance of a representative reference group of Dutch primary schools. Furthermore, school reports compare teachers' scores with those of the principals, and classroom reports compare teachers' scores with those of pupils (Hendriks & Bosker, 2003).

Theoretical framework

A framework developed by Visscher (2002) for studying School Performance Feedback Systems (SPFSs) was used as the theoretical framework for this study. According to Coe and Visscher (2002a) 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). ZEBO is one of those systems which provide schools with confidential information, in the form of school and classroom reports, on their performance and functioning as a basis for school self-evaluation. ZEBO is, however, a school internal system. Based on Visscher’s framework a new framework was devised for this study.

According to the framework, see Figure 1, ZEBO use (D) is influenced by the characteristics of ZEBO (as perceived by the users of ZEBO) (A), implementation process features (B), and school organisational characteristics (C). Furthermore ZEBO use (D) is, in turn, expected to have certain effects (E) on, for example, pupil achievement and several other prerequisites for school improvement, such as the instructional strategies of the teachers and the functioning of the principal.

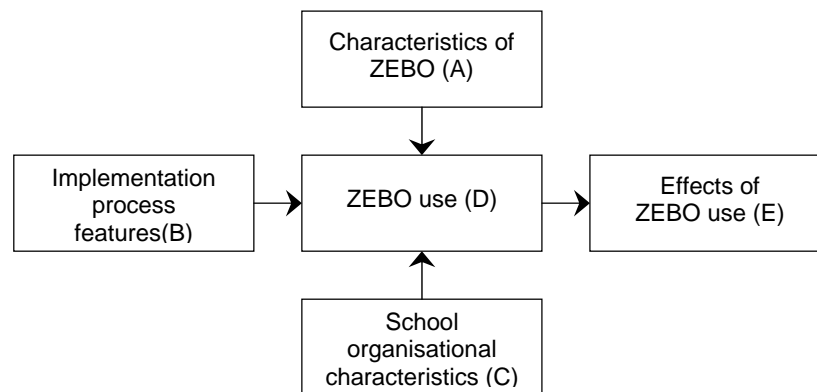


Figure 1 The use, effects and mediators of the use of ZEBO (Based on: Visscher, 2002)

For this study, the variables in the five groups of factors in the framework have been specified. *ZEBO use (D)* assumes the inclusion of the study and discussion of the ZEBO output. That information is meant to provide practitioners with new insights. The ZEBO output may draw attention to certain problems within the school, and a school team or an individual staff member may decide to attempt to devise solutions. Finally, the use of output generated by ZEBO may lead to (policy) measures taken at school level and at classroom level, by the whole school or by individual teachers, with the goal of school improvement.

Characteristics of ZEBO (A) concern characteristics of ZEBO use perceived by its users, such as the relevancy of the output. ZEBO characteristics may vary with regard to how relevant, up-to-date, accurate, and fitting the needs of the users, the ZEBO output is

perceived to be, by the users of ZEBO. The user-friendliness of the system is another characteristic thought to be an important factor influencing ZEBO use. This refers to data entry, altering input, information retrieval, and interpretation of results. Finally, the amount of time and effort perceived to be required to use ZEBO may influence ZEBO use.

Implementation process features (B) may influence the use of ZEBO. These include user support and training; whether or not schools are satisfied with the amount of training and support they have received; support of the principal; the amount of pressure and support received; and whether the purpose of the self-evaluation was clear to all school staff.

It is hypothesized that *school organizational characteristics (C)* may also influence ZEBO use. These include: the innovation attitude of school staff which depends largely on the perceived advantages and disadvantages of ZEBO; time and resources for innovation activities; and a school's innovation capacity. Finally, the score on the ZEBO scales (ZEBO score) is considered relevant for use of ZEBO. The score on ZEBO may influence whether a school chooses to use the ZEBO output in order to improve school performance. Relatively low scores on ZEBO scales combined with a pressure strategy may increase the motivation to improve performance in a school by using the ZEBO output.

ZEBO use (D) is, in turn, expected to have certain *Effects (E)*. If school staff use the ZEBO output, this may have unintended as well as intended effects. ZEBO was designed to monitor and improve school performance. Next to the effect of ZEBO use on pupil achievement, this study also explored the degree to which specific prerequisites for improved pupil achievement could be observed. Whether or not ZEBO use had an effect on the following variables was investigated: the amount of consultation on the schools' functioning, the principal's functioning, the professional development of school staff, the achievement orientation of the school, team cohesion, pupil care, teacher behaviour in the classroom, the amount of evaluation conducted within the school, and adaptive instruction. Finally, the use of ZEBO may have effects, other than those intended, such as negative impact on teachers as a result of the workload ZEBO imposes on them.

Method

The use of ZEBO was studied in 79 primary schools in the Dutch Twente region. Participation by schools was voluntary. All primary schools in this region were asked to participate. The number of public schools is larger and the number of Protestant schools is smaller in the convenience sample than in the population. The number of Catholic schools however, is representative of the population. The schools in the sample had a smaller average school size ($F=10.61$; $p=0.01$), but the sample was representative in terms of the composition of the pupil population of schools (in terms of pupil socio-economic status and ethnicity) ($F=0.26$; $p=0.61$).

Schools administered ZEB0 three times, in 2003, 2004, and in 2006. To answer the research questions, a questionnaire (Evaluation of ZEB0 Questionnaire) was constructed based on the framework described above. For almost all questionnaire items, a statement format with a four 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 questionnaire was designed to study the five groups of factors shown in Figure 1. The number of items for each group of factors varied as follows:

- § (A) Characteristics of ZEB0 as perceived by its users (10 items);
- § (B) Implementation process features (8 items);
- § (C) School organisational characteristics (20 items);
- § (D) ZEB0 use: divided into conceptual use (4 items) and instrumental use (5 items)+
- § (E) Effects of ZEB0 (10 items).

After each of the three administrations of ZEB0, teachers and principals completed the Evaluation of ZEB0 Questionnaire. Each time, interviews were then conducted with both teachers and principals to obtain: more detailed information on the use of ZEB0; the factors influencing the use of ZEB0; and to validate the results of the Evaluation of ZEB0 Questionnaire. Each time (2003, 2004, and 2006) the schools were divided into three groups based on the results of the nine ZEB0 use (D) variables in the questionnaire and respondents from each group were interviewed:

- § LoSE (Low Self-Evaluation): schools which did not use the ZEB0 output;
- § AvSE (Average Self-Evaluation): schools which made average use of the ZEB0 output;
- § HiSE (High Self-Evaluation): schools which made intensive use of the ZEB0 output.

In 2003, 31 respondents were interviewed and 25 respondents in 2004. In 2006, due to time restrictions and organisational problems, only 2 respondents were interviewed.

To study the effect of ZEB0 use on pupil achievement, two cohorts of pupils were followed by means of two standardized achievement tests (spelling and mathematics) over the five year period.

The first question to be answered was “How and to what extent do schools ZEB0?” In analysing the Evaluation of ZEB0 Questionnaires, a distinction was made between instrumental use and conceptual use of ZEB0 (Based on Weiss, 1998a). Rossi, Freeman and Lipsey (1999) differentiate between these two uses of evaluation findings, and define instrumental use as the direct use of evaluation findings decisions and actions based on the evaluation. Conceptual use refers to the indirect use of feedback which may influence thinking about issues in a more general way, and in the longer term, may have an impact on the users’ actions.

The data obtained through the interviews were analysed using the Atlas/ti program. Answering Research Question one “How and to what extent do schools use ZEBO?” led to a description of perceived ZEBO use within schools.

Data provided by the questionnaires were also used to answer Research Question two “What are the effects of ZEBO use?” The data on the perceived effects (E) were calculated for each moment of measurement by means of frequencies. In addition, multilevel analyses with repeated measures were conducted to investigate the effect of ZEBO use on pupil achievement.

To answer Research Question three “Which factors influence the use of ZEBO?” data analyses started with the analysis of correlations between the ZEBO use (D) variables and the independent variables (A) (B) & (C) in Figure 1, for both the principals’ and the teachers’ data. Those variables which correlated significantly and strongly with the use of ZEBO were entered into regression and multilevel 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 teachers’ data. The interview results were analysed with the help of the Atlas program.

Q.1 How and to what extent do schools use ZEBO?

The quantitative analyses of the Evaluation of ZEBO Questionnaire data and the qualitative analyses (by means of the Atlas program) of the interview data show that in 2003 both conceptual and instrumental ZEBO uses (D) were limited, especially use by teachers. A possible explanation of the limited conceptual use of ZEBO may be that some teachers never received the ZEBO output, as became clear from the interviews. Many teachers (44%) indicated that they did not study, or barely studied the ZEBO output. In most cases, the principals controlled the ZEBO administration, and the ZEBO output, and therefore controlled the dissemination of the output.

For instrumental use of ZEBO output the results are similar to those for conceptual use. Respondents from 12 of 50 (24%) schools, who worked with ZEBO, used the results instrumentally in 2003. These respondents indicated that the results were discussed. Moreover, respondents from these schools mentioned that actions were taken to improve the quality of education. Measures taken included: stimulating independent learning; increased instructional differentiation; implementation of classroom consultation; and more frequent evaluation of pupils.

The conceptual use of ZEBO was still limited in 2004. At that time fewer principals reported that the ZEBO output led to new insights. It is suggested that as schools were using ZEBO for the second time, perhaps principals knew what to expect. A higher number of teachers reported new insights as a result of using ZEBO. This may be due to the fact that in 2003 almost half of the teachers had not seen the ZEBO output, whereas

in 2004, 141 of the 236 teachers (61%) who completed the questionnaire indicated that they had studied the ZEBO output.

Instrumental use was also still limited in 2004. Thirteen schools (26%) used the ZEBO output instrumentally at that time. Teachers from these schools reported that the ZEBO output provided a common starting point for discussion. Extensive discussions may help school staff to arrive at a deeper understanding of the ZEBO output, and of the implications for their work. The results of the second evaluation of ZEBO use (2004), showed that the teachers and principals from the schools which used the ZEBO output instrumentally, indicated that the school team took measures, based on the ZEBO output, to improve the quality of education.

In 2006, (most) schools used ZEBO for the third time. School staff used the results more than they had in 2003 and in 2004. It must be noted here, that only 43 schools administrated ZEBO in 2006, compared to 64 schools in 2003 and 58 schools in 2004. This decline in ZEBO administration was mainly due to the fact that twelve schools chose to postpone administering ZEBO until 2007. In 2006, staff in 32% of the schools (14 schools) used the ZEBO output instrumentally. Teachers indicated in the questionnaires and interviews that they accommodated to pupil differences more, made their education more adaptive, and brought more clarity to their lessons. Principals used the output for making changes at school level, for example, for writing the school plan, for school policy development, and for conducting performance interviews.

Q.2 What Are the Effects of ZEBO Use?

In the multilevel analyses with repeated measures, conducted to study the effect of ZEBO use (D) on pupil achievement (E), no significant relationships were found between the use of ZEBO and pupil achievement. The results also indicate that the use of ZEBO had no negative effects.

ZEBO use was found to have some other effects (E) in the limited number of schools which were using the ZEBO output. The Evaluation ZEBO Questionnaire results indicated that ZEBO use led, in these schools, to an improvement in consultation and communication. In addition, quality care became an important subject in team meetings, the mission of the school was discussed more often, and the use of ZEBO led to more discussions on how to improve the school's functioning. Moreover, didactic behaviour, achievement orientation, and professional development also improved as a result of ZEBO use in 2006. To summarise: although the effects of ZEBO use are limited, to date, the results of this study indicate that, in schools which used the ZEBO output, some effects were found on important prerequisites for school improvement.

Q.3 Which factors influence the use of ZEBO?

The results of the regression analyses, multilevel analyses, and the interviews show that the characteristics of ZEBO as perceived by its users (A), implementation process features (B), and school organisational characteristics (C) influenced ZEBO use (D).

Characteristics of ZEBO (A)

Firstly, three characteristics of ZEBO (A) were found to influence ZEBO use (D). Schools, which were more likely to use the ZEBO output to make improvements, were found to be those in which the following characteristics of ZEBO were judged more positively:

- § Fit of output with user needs;
- § The ease of data input;
- § Time requirement of use (these users reported that the use of ZEBO was not time-consuming).

Characteristics of ZEBO were found to influence ZEBO use in the first two evaluations only (in 2003 and 2004). In 2006, these characteristics of ZEBO appear not to have had any further influence on the use of ZEBO. In general, schools which judged the ZEBO characteristics more negatively chose not to continue using ZEBO. At the end of 2003 and 2004 eleven schools chose to stop using ZEBO and continued with another school self-evaluation instrument.

Implementation Process Features (B)

Secondly, three Implementation process features (B) were found to influence ZEBO use (D). The findings suggest that users who used ZEBO more intensively rated the following factors more positively. The degree to which:

- § The goal ZEBO is clear. In 2003, at the start of this study, the goal of ZEBO was not clear to several teachers, and in 2006, the goal was not clear to several principals. The latter was probably due to the fact that several schools had new principals at that time, who may not have been familiar with ZEBO;
- § School staff feels they are trained sufficiently in the use of ZEBO. Satisfaction with the training received only played a role in 2004 (when schools used ZEBO for the second time). In 2003 most schools did not use the ZEBO output. In 2004, schools started to use the ZEBO output and they probably realised then that they would benefit from some training in the use of ZEBO. In 2006, teachers and principals indicated that they did not need any further training;
- § The principal encourages the use of the self-evaluation instrument. Findings showed that the principal played a consistently important role in the use of the self-evaluation output in 2003, and in 2004. In 2006, however, this variable did not explain variance in the use of ZEBO, but in that year, 86% of the teachers indicated that the principal encouraged the use of ZEBO.

In addition, the interview results indicate that the degree to which the principal takes responsibility for the ZEBO output influences the use of ZEBO. The interview results indicate that it is important that the principal distributes the results, explains them where necessary, and deliberates with school staff on how best to use the results to improve the school functioning.

School Organizational Characteristics (C)

Thirdly, school organizational characteristics (C) were found to influence the use of ZEBO. Schools which used the ZEBO output more were those schools in which a positive innovation attitude already existed. Which aspects of the innovation attitude of staff explained variance in ZEBO use were also investigated? Three items included in the innovation attitude scale were found to influence ZEBO use, namely the degree to which:

- § School staff believes that the use of ZEBO will lead to quality improvement;
- § School staff are not afraid (in advance) of the changes ZEBO use could cause;
- § Teachers feel they are able to influence the measures taken based on ZEBO.

Another school organizational characteristic consistently found to be influencing the use of ZEBO is the innovation capacity of schools. Schools with a high innovation capacity were found to be more likely than those with lower innovation capacity to use the ZEBO output. Which items included in the innovation capacity scale were responsible for variance in ZEBO use was also explored. The following six items from the scale were found to influence the use of the ZEBO, the degree to which:

- § The school monitors the quality of education on a regular basis;
- § School staff experiments with how to improve their education;
- § The principal encourages professional development of school staff;
- § Teachers exchange information on their functioning;
- § The decision to participate in the ZEBO project was made by the whole school;
- § The school team is not cohesive. Weaker team cohesion led in this study to more ZEBO use in some schools but only in 2006. The reason for this is not clear. It is possible that the ZEBO output is used to improve team cohesion in these schools.

In addition, schools which had extra time and resources made more use of ZEBO output. The interview results show, for example that in several schools teachers indicated that they needed help with the ZEBO administration (specifically with the ZEBO pupil questionnaire). The factor dealing with time and resources was important in 2004 and remained important in 2006. There are time and money costs involved in ZEBO use, for example, to discuss the results extensively, to take professional development courses, if needed, and to implement new teaching methods.

Another school organisational characteristic which was found to influence the use of the ZEBO output was the ZEBO score. If a school scores on or above average, on the process indicators measured by ZEBO, there is probably no need, or a reduced need to use the ZEBO output, since this means that the school performs at or better than the average Dutch school. To take this aspect into account, a variable representing the number of scales on which the school scored on or above average, was calculated (the ZEBO score). The ZEBO score only played a role in 2003 but no longer in 2004 and 2006. A higher number of schools scored on or above average on more ZEBO scales during the second and third evaluation, leading to less variance.

The interview results indicate that two additional factors might have influenced the use of ZEBO:

- § The degree to which the self-evaluation output is congruent with the expectations of the principal. If the self-evaluation results are not congruent with the expectations of the principal, the principal may reject the results, which may be the end of the self-evaluation process;
- § The degree to which an open school climate exists. School staff must be receptive to criticism, and it must be possible to discuss issues. The two principals interviewed in 2006 hypothesized that this variable could play a role in the use of self-evaluation output. It seems plausible that schools in which school staff is unable to accept criticism or discuss issues, the use of self-evaluation results will be difficult, if not impossible.

Conclusion and discussion

In the concluding chapter, firstly a reflection on the research design is presented. One of the weaknesses of this study is for example that the sample was not representative for school denomination and school size, which may limit the generalizability of the research findings.

Secondly, the practical implications of the results of this study for schools are presented. The findings of this investigation give rise to several practical implications for the successful implementation and use of SPFSs in schools. These serve to highlight the importance of such factors as: the need for the entire school staff to be involved in the decision-making processes regarding, for example, which instrument to choose; the need for the goal of the self-evaluation process to be clear to school staff; and the need for the principal to encourage school staff to use the self-evaluation instrument.

Thirdly, practical implications for school self-evaluation designers include, among others, the suggestion that the use of self-evaluation results may be increased by the provision of guidelines on how to use the self-evaluation results, and examples of actions a school may undertake to improve its functioning.

Fourthly, the results of this study also have implications for theories of school performance feedback. The SPFS framework has proven to be very suitable for studying the use of an instrument such as ZEBO. However, it was not possible to explain all the variance in ZEBO use by means of the framework. This may be due, for example, to the quality of the instruments used in this study, but it may also imply that the framework does not cover all the variables that may play a role in the use of performance feedback. Some variables are suggested to add to the framework, including the degree to which the feedback include cues on how to improve one's performance and the degree to which other sources of relevant information are seriously competing for the attention of the users. Finally, it is suggested that characteristics of the self-evaluation users (such as their motivation, locus of control, and feelings of uncertainty) may influence the use of feedback.

The thesis ends with some suggestions for further research. This study provides more insight into the use and effects of school self-evaluation systems and performance feedback, and into the factors influencing the use of performance feedback. However, it also raises questions in new areas, including the role of (other) data use for school improvement. Outside of school self-evaluation results, schools have other data available to them to assess their functioning and improve their performance, such as student assessment data. Questions such as “What practices and support may be provided to, and by, school staff in improving effective and enhanced data use?” require further research. Moreover, much is still unknown about the effects and mediators of school performance feedback. Since quality care and the development of school evaluation instruments receive much attention around the globe, the justifications for using self-evaluation instruments are plausible. As thousands of schools are using school self-evaluation instruments, further research into the use of these systems and into the conditions promoting effective self-evaluation use is desirable.

Samenvatting

Inleiding en onderzoeksvragen

Scholen over de hele wereld maken gebruik van zelfevaluatie-systemen, maar er is weinig bekend over hoe scholen deze systemen gebruiken, welke effecten gebruik heeft en wat de voorwaarden voor succesvolle zelfevaluatie zijn (Coe & Visscher, 2002a). Om die reden is er in 2001 een onderzoek gestart naar het gebruik en de effecten van een zelfevaluatie-instrument dat gebruikt wordt in Nederland: ZEBO (ZelfEvaluatie BasisOnderwijs). Drie vragen lagen ten grondslag aan het onderzoek:

1. In welke mate en hoe gebruiken scholen ZEBO?
2. Wat zijn de effecten van het gebruik van ZEBO?
3. Welke factoren zijn van invloed op het gebruik van ZEBO?

Zelfevaluatie werd in deze studie gedefinieerd als een door de school zelf geïnitieerde procedure voor het systematisch verzamelen van informatie om het functioneren van de school en de mate waarin de school de door haar gestelde doelen bereikt heeft vast te stellen, met als doel het beleid van de school te ondersteunen en het leren te stimuleren. Het uiteindelijke doel van zelfevaluatie is schoolverbetering.

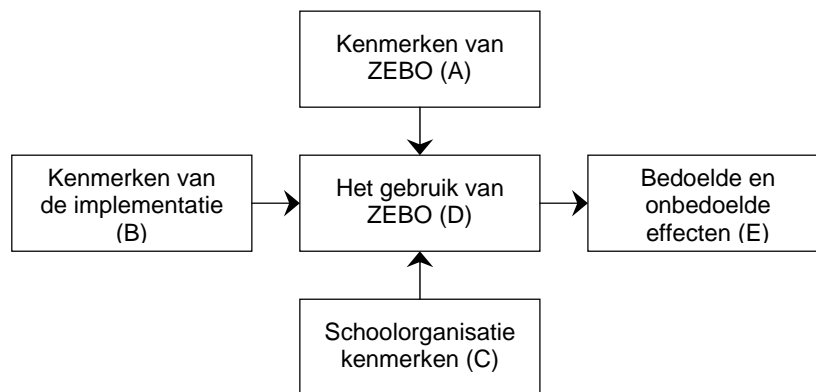
ZEBO (ZelfEvaluatie BasisOnderwijs) is een zelfevaluatie-instrument voor het Nederlandse basisonderwijs. Het is een instrument voor het meten van aspecten van processen op klasniveau en op op basisscholen, met schooleffectiviteit en instructie-effectiviteit als conceptueel kader. ZEBO bestaat uit een reeks vragenlijsten voor leerlingen, leerkrachten en de schoolleiding die met behulp van de computer worden afgenomen en verwerkt. De schoolleiding wordt gevraagd een oordeel te geven over de organisatie op schoolniveau en over het functioneren van de schoolleiding: bijvoorbeeld vragen over formeel overleg, samenwerking, onderwijskundig leiderschap, professionalisering, doelstellingen en verwachtingen, teamcohesie, functioneren van de schoolleiding, werkbelasting en planmatig handelen. De leerkrachten beantwoorden vragen over het functioneren van de school als organisatie, het functioneren van de schoolleiding en het onderwijs in de groep. Bij deze laatste categorie vragen gaat het deels om dezelfde onderwerpen als bij de leerlingen, zodat antwoorden van leraren en leerlingen vergeleken kunnen worden. Leerlingen wordt gevraagd om een oordeel te geven over het onderwijs in hun groep. Ze beantwoorden daarbij vragen over de leerkracht, medeleerlingen, prestatiedruk, didactisch handelen, hun activiteiten in de klas (krijgt men adaptief les), het werk (leertijd) en de werksfeer in de klas.

Na het invullen van de vragenlijsten kan met ZEBO een rapportage gegenereerd worden. Een belangrijk kenmerk van de feedback die zo verkregen wordt is dat de scores op klasniveau en schoolniveau vergeleken kunnen worden met het nationale gemiddelde van Nederlandse basisscholen. Ook kunnen de scores van de schoolleiding vergeleken worden met de scores van de leerkrachten, en kunnen de scores van de leerkrachten vergeleken worden met de scores van de leerlingen (Hendriks & Bosker, 2003).

Theoretisch kader

ZEBO kan gekenmerkt worden als een (intern) School Performance Feedback Systeem (SPFS). Dit zijn externe systemen die scholen kunnen gebruiken om betrouwbare informatie te verkrijgen over hun functioneren, als een basis voor de evaluatie van hun presteren en functioneren. “School improvement” is het belangrijkste doel van een SPFS (Visscher & Coe, 2002). Visscher (2002) heeft een raamwerk ontwikkeld om het gebruik van SPFS-en te onderzoeken. Dit raamwerk is gebruikt als basis voor het onderzoek naar ZEBO.

In figuur 1 is te zien welke factoren van belang geacht worden bij het *gebruik van ZEBO* (D). De *kenmerken van ZEBO* (A) kunnen het gebruik van ZEBO beïnvloeden, evenals de *kenmerken van het implementatieproces* (B) en de *schoolorganisatiekenmerken* (C). ZEBO gebruik kan zowel bedoelde als onbedoelde effecten (E) hebben.



Figuur 1 Het gebruik, de effecten en de factoren die het gebruik van ZEBO mogelijk beïnvloeden (gebaseerd op Visscher, 2002)

Bij het gebruik van de ZEBO resultaten (D) gaat het mogelijk om het bestuderen en bespreken van de ZEBO-resultaten; de ZEBO-resultaten kunnen tot nieuwe inzichten bij gebruikers leiden. Zo kunnen de ZEBO-resultaten op bepaalde problemen binnen de school wijzen en het schoolteam of een individuele leerkracht of directeur kan oplossingen voor deze problemen bedenken. Bij de ZEBO-kenmerken (A) gaat het om de kenmerken van ZEBO, zoals beoordeeld door de gebruikers. De mate waarin gebruikers vinden dat de informatie uit ZEBO relevant, up-to-date, juist is en de mate waarin de informatie aansluit bij de behoeften van de school, kan variëren. Ook wordt verondersteld dat het gebruik van ZEBO zal toenemen als het niet te moeilijk is om data

in te voeren, wijzigingen aan te brengen en informatie uit ZEBO te halen. De informatie moet ook begrijpbaar zijn en werken met ZEBO moet niet te veel tijd kosten.

De manier waarop ZEBO wordt geïmplementeerd (B) zal het gebruik ervan waarschijnlijk eveneens beïnvloeden. Stimulering van het gebruik door de schoolleider, training en ondersteuning, druk die men voelt om ZEBO te gebruiken en duidelijkheid omtrent het doel van ZEBO, zijn allen factoren die hierbij een belangrijke rol kunnen spelen.

Ten derde beïnvloeden schoolorganisatiekenmerken (C) mogelijk het gebruik van ZEBO. Scholen moeten het nut inzien van veranderingen en ze dienen te beschikken over voldoende innovatiecapaciteit. Daarnaast spelen de beschikbaarheid van middelen (zoals de beschikbaarheid van geld en personeel om met ZEBO te werken) en de scores op ZEBO mogelijk een rol bij het gebruik van ZEBO. Met betrekking tot de ZEBO scores is het waarschijnlijk dat als scholen al hoog scoren op ZEBO (in vergelijking met het landelijk gemiddelde) scholen weinig druk zullen voelen om hun onderwijs te verbeteren.

Het gebruik van ZEBO, zoals hierboven beschreven, kan bedoelde en onbedoelde *effecten* (E) hebben. Het doel van het gebruik van ZEBO is de verbetering van de prestaties van de school. Het eerste effect dat dus onderzocht moet worden is een effect op de leerprestaties. ZEBO geeft scholen informatie op schoolniveau en op klasniveau. Het gebruik van ZEBO kan dus ook effecten hebben op school- en op klasniveau. Het is onderzocht of ZEBO een effect heeft gehad op de volgende variabelen: de mate van overleg over het functioneren van de school en de kwaliteit van het onderwijs, het functioneren van de schoolleider, professionele ontwikkeling, prestatie oriëntatie, collegialiteit, leerlingzorg, het lesgeven, de evaluatie van leerlingprestaties en adaptief onderwijs. Tot slot kan het gebruik van ZEBO ook onbedoelde negatieve effecten hebben, zoals een negatief effect op leerkrachten door de toenemende werkdruk als gevolg van het gebruik van ZEBO.

Methode

Het gebruik van ZEBO is onderzocht in 79 basisscholen in Twente. De deelname aan het onderzoek was vrijwillig. Alle basisscholen in de regio is gevraagd om deel te nemen aan het onderzoek. Het aantal openbare scholen in de gelegenhedssteekproef is groter dan in de populatie en het aantal protestante scholen is kleiner dan in de populatie. Het aantal katholieke scholen is representatief voor de populatie. De scholen in de steekproef zijn gemiddeld kleiner ($F=10.61$; $p=0.01$), maar de steekproef is wel weer representatief in termen van de sociaal-economische status van leerlingen en hun etniciteit ($F=0.26$; $p=0.61$).

Scholen hebben ZEBO drie keer gebruikt in een periode van 5 jaar (in 2003, 2004 en 2006). Om het gebruik en de effecten van ZEBO te evalueren is er een schoolleider- en een leerkrachtvragenlijst ontwikkeld op basis van het raamwerk ontwikkeld door Visscher (2002) en op basis van schooleffectiviteitsliteratuur: de vragenlijst "Evaluatie

van het gebruik van ZEBO". Deze vragenlijst bestaat uit een aantal gesloten vragen en een aantal open vragen. De gesloten vragen hebben een vierpuntschaal (van 1, oftewel 'helemaal mee eens' tot 4 'helemaal mee oneens'). De vragenlijst bestaat uit verschillende vragen om de factoren in figuur 1 te meten:

- § (A) Kenmerken van ZEBO (10 items);
- § (B) Kenmerken van de implementatie (8 items);
- § (C) Schoolorganisatiekenmerken (20 items);
- § (D) Het gebruik van ZEBO: verdeeld in conceptueel gebruik (4 items) en instrumenteel gebruik (5 items);
- § (E) Bedoelde en onbedoelde effecten van ZEBO (10 items).

Leerkrachten en schoolleiders hebben deze vragenlijsten drie keer ingevuld (telkens 3 tot 6 maanden na de afname van ZEBO). Om het gebruik van ZEBO verder te onderzoeken werden tevens elk jaar interviews gehouden met leerkrachten en schoolleiders van scholen die met ZEBO gewerkt hadden. Op basis van de resultaten van de vragenlijsten werden de scholen in drie groepen gebruikers verdeeld:

- § Een groep die geen gebruik had gemaakt van de ZEBO-resultaten (LoSE: Low Self Evaluation);
- § Een groep die gemiddeld gebruik had gemaakt van de resultaten (AvSE: Average Self Evaluation);
- § Een groep die veel gebruik had gemaakt van de resultaten (HiSE: High Self Evaluation).

In 2003 zijn 31 respondenten geïnterviewd uit 4 LoSE-scholen, 5 AvSE-scholen en 2 HiSE-scholen. In 2004 zijn 25 respondenten geïnterviewd uit 3 LoSE-scholen, 3 AvSE-scholen en 3 HiSE-scholen. Door organisatieproblemen zijn in 2006 maar twee personen geïnterviewd.

Ten slotte: om te na te gaan of een verband bestaat tussen ZEBO-gebruik en leerprestaties, werden de scores op de toetsen spelling en rekenen van het leerlingvolgsysteem van Cito gedurende 5 jaar verzameld. Leerlingen werden daarbij gevolgd van groep 3/4 tot groep 7/8. Ook werd van deze leerlingen achtergrondinformatie verzameld zoals hun geslacht, sociaal- economische status, leeftijd, geschat IQ en thuistaal.

De eerste onderzoeksvraag was: "Hoe en in welke mate gebruiken scholen ZEBO?". Bij het analyseren van de "evaluatie van het gebruik van ZEBO"-vragenlijsten is onderscheid gemaakt tussen conceptueel gebruik en instrumenteel gebruik (gebaseerd op Weiss, 1998a). Rossi, Freeman en Lipsey (1999) definiëren instrumenteel gebruik als het direct gebruiken van de evaluatieresultaten. De resultaten worden geanalyseerd en er worden maatregelen genomen op basis van de resultaten. Bij conceptueel gebruik van de resultaten gaat het erom dat de resultaten het denken van de mensen beïnvloeden. De data

die verkregen waren met behulp van de interviews is geanalyseerd met behulp van het programma Atlas/ti. Het beantwoorden van de eerste onderzoeksvraag heeft geleid tot een beschrijving van het gebruik van ZEBO door scholen.

De data verkregen met behulp van de vragenlijsten zijn ook gebruikt voor het beantwoorden van de tweede onderzoeksvraag: “Wat zijn de effecten van het gebruik van ZEBO?” Op basis van de data met betrekking tot de gepercipieerde effecten van ZEBO werden frequenties berekend. Daarnaast zijn meerniveau analyses uitgevoerd om het effect van ZEBO op leerprestaties te onderzoeken.

Om onderzoeksvraag 3 (“Welke factoren zijn van invloed op het gebruik van ZEBO?”) te beantwoorden zijn correlaties berekend tussen het gebruik van ZEBO (D) en de onafhankelijke variabelen in blok (A), (B) en (C) in figuur 1, zowel voor schoolleiders als voor leerkrachten. De variabelen die significant correleren met het gebruik van ZEBO zijn vervolgens in regressie- en meerniveau analyses ingevoerd. Regressie-analyses zijn uitgevoerd op de data van de schoolleiders. Omdat de leerkrachten data een geneste structuur hebben (leerkrachten binnen scholen) zijn meerniveau analyses uitgevoerd op de data van de leerkrachten. De interviewdata zijn geanalyseerd met behulp van het Atlas/ti-programma.

Vraag 1: Hoe en in welke mate gebruiken scholen ZEBO?

Frequentieanalyses en de analyse van de interviews met behulp van het programma Atlas laten zien dat het gebruik van ZEBO na de eerste afname van ZEBO beperkt is, vooral het gebruik door leerkrachten. Een mogelijke verklaring hiervoor is dat een aantal leerkrachten de resultaten nooit gezien heeft. Veel leerkrachten (44%) gaven aan de ZEBO-resultaten niet of nauwelijks bestudeerd te hebben. Schoolleiders beheerden het programma meestal en hadden de uitslag dus wel tot hun beschikking.

De resultaten m.b.t. het instrumenteel gebruik van ZEBO laten hetzelfde beeld zien. Respondenten van 12 van de 50 scholen (24%) hebben de resultaten instrumenteel gebruikt. Deze respondenten gaven aan dat de resultaten besproken waren. Op deze scholen zijn ook verschillende maatregelen genomen om de kwaliteit van het onderwijs te verbeteren. Maatregelen die genoemd werden, waren het verbeteren van de communicatie binnen de school, het stimuleren van zelfstandig leren (bijvoorbeeld door de invoer van blokuren), meer differentiatie, het invoeren van klassenbezoek en frequenter evalueren en testen van leerlingen.

De evaluatie van het tweede gebruik van ZEBO (2004) laat zien dat het gebruik nog steeds beperkt is. Minder schoolleiders gaven aan dat het gebruik van ZEBO tot nieuwe inzichten heeft geleid. Het is mogelijk dat schoolleiders wisten wat ze konden verwachten, doordat scholen ZEBO voor de tweede keer gebruikten. In vergelijking met 2003 gaven in 2004 meer leerkrachten aan dat de resultaten tot nieuwe inzichten hebben geleid. Deze toename wordt waarschijnlijk veroorzaakt doordat in 2003 bijna de helft

van de leerkrachten aangaf de ZEBO-resultaten niet gezien te hebben, terwijl in 2004 141 van de 236 leerkrachten (61%) aangaven de resultaten gezien te hebben. Instrumenteel gebruik is ook in 2004 nog steeds beperkt. Dertien scholen (26%) hebben de resultaten instrumenteel gebruikt in 2004. Leerkrachten van deze scholen gaven aan dat de ZEBO-resultaten hen een startpunt voor discussie boden. Ook in 2004 gaven schoolleiders en leerkrachten van de scholen die de ZEBO-resultaten instrumenteel gebruikt hadden aan dat ze op basis van deze resultaten maatregelen hadden genomen om de kwaliteit van hun onderwijs te verbeteren.

In 2006 hebben (de meeste) scholen ZEBO voor de derde keer gebruikt. Uit de derde evaluatie blijkt dat het gebruik van ZEBO in 2006 wederom is toegenomen. Wel moet hierbij vermeld worden dat in 2006 43 scholen ZEBO hebben afgenomen, tegenover 64 scholen in 2003 en 58 scholen in 2004. Deze afname werd voornamelijk veroorzaakt door 12 scholen die ervoor kozen om ZEBO pas in 2007 weer af te nemen. In 2006 heeft 32% van de scholen (14 scholen) ZEBO instrumenteel gebruikt. Leerkrachten gaven in de vragenlijsten en interviews aan dat ze hun lessen meer probeerden aan te passen aan de verschillen tussen leerlingen en dat ze probeerden meer adaptief en duidelijker les te geven. Schoolleiders hebben de ZEBO-resultaten vooral gebruikt voor het doorvoeren van veranderingen op schoolniveau, bijvoorbeeld voor het schrijven van het schoolplan, de ontwikkeling van beleid en voor het voeren van functioneringsgesprekken.

Vraag 2: Wat zijn de effecten van het gebruik van ZEBO?

Met behulp van meerniveau analyses is onderzocht of het gebruik van ZEBO een effect had op de leerprestaties van leerlingen. Uit de resultaten van deze analyses bleek dat het gebruik van ZEBO (D) geen effect heeft gehad op de leerprestaties van leerlingen (E). De resultaten van de vragenlijsten en interviews laten verder zien dat het gebruik van ZEBO ook geen negatieve effecten heeft gehad.

Echter, de resultaten laten zien dat het gebruik van ZEBO wel effect gehad heeft op andere belangrijke voorwaarden voor schoolverbetering. Zo heeft het gebruik van ZEBO geleid tot meer overleg over het functioneren van de school en over de kwaliteit van het onderwijs. Ook werd de missie van de school vaker besproken en werd meer gesproken over de wijze waarop het onderwijs verbeterd kon worden. Andere effecten die genoemd werden waren: meer aandacht voor de professionele ontwikkeling van het team, een toename van de prestatieoriëntatie, en betere leskwaliteit. Samenvattend: hoewel het gebruik van ZEBO door de jaren heen beperkt is gebleven, laten de resultaten van dit onderzoek zien dat het gebruik wel enige effecten heeft gehad in de scholen die de ZEBO-resultaten daadwerkelijk gebruikt hebben.

Vraag 3: welke factoren zijn van invloed op het gebruik van ZEBO?

De resultaten van de regressieanalyses, meerniveau analyses en de interviews laten zien dat een aantal kenmerken van ZEBO (A), het implementatieproces (B) en de schoolorganisatie (C) het gebruik van ZEBO (D) beïnvloeden.

Kenmerken van ZEBO (A)

Drie kenmerken van ZEBO, zoals beoordeeld door de gebruikers hebben het gebruik van ZEBO beïnvloed (D):

- § De mate waarin de informatie uit ZEBO aansluit bij de behoeften van de school;
- § De mate waarin het makkelijk is om gegevens in ZEBO in te voeren;
- § De hoeveelheid tijd die het werken met ZEBO kost.

De kenmerken van ZEBO beïnvloedden het gebruik van ZEBO alleen in 2003 en 2004. In 2006 lijken deze kenmerken geen invloed meer te hebben op het gebruik van ZEBO. In het algemeen zijn de scholen die de kenmerken van ZEBO negatief beoordeelden gestopt met het gebruiken van ZEBO. In 2003 en 2004 zijn 11 scholen gestopt met ZEBO en verder gegaan met een ander zelfevaluatie-instrument.

Kenmerken van de implementatie (B)

Drie implementatiekenmerken hebben het gebruik van ZEBO (D) beïnvloed. De resultaten laten zien dat intensievere ZEBO-gebruikers de volgende implementatiekenmerken positiever beoordeelden. De mate waarin:

- § Het doel van ZEBO duidelijk is. Aan het begin van het onderzoek in 2003 was het doel van het gebruik van ZEBO voor veel leerkrachten niet duidelijk. In 2006 was het doel niet duidelijk voor een aantal schoolleiders. Dit laatste werd waarschijnlijk veroorzaakt door het feit dat in die periode een aantal scholen nieuwe schoolleiders had gekregen die hoogstwaarschijnlijk niet bekend waren met ZEBO;
- § Het schoolteam meent voldoende getraind te zijn in het gebruik van ZEBO. Deze variabele speelde alleen een rol in 2004 (bij de tweede keer dat scholen ZEBO gebruikten). In 2003 hebben de meeste scholen ZEBO niet gebruikt. In 2004 begonnen de scholen langzaam gebruik te maken van ZEBO en leerkrachten en schoolleiders realiseerden zich op dat moment waarschijnlijk dat ze training in het gebruik van ZEBO nodig hadden. In 2006 was deze variabele niet meer van invloed. Schoolleiders en leerkrachten gaven toen aan geen training meer nodig te hebben;
- § De schoolleider het gebruik van ZEBO stimuleert. De schoolleider speelde een belangrijke rol in het gebruik van ZEBO in 2003 en 2004. In 2006 verklaart deze variabele geen variantie meer in het gebruik van ZEBO. Nadere analyse maakt duidelijk dat dit veroorzaakt wordt door het feit dat in 2006 86% van de leerkrachten aangeeft dat de schoolleider het gebruik van ZEBO stimuleert.

Verder laten de interviewresultaten zien dat de mate waarin de schoolleider verantwoordelijkheid neemt voor de ZEBO-resultaten het gebruik van ZEBO ook kan beïnvloeden. De interviewresultaten maken duidelijk dat de schoolleider de resultaten moet verspreiden, de resultaten moet uitleggen waar dit nodig is en met het team moet overleggen hoe de resultaten het beste gebruikt kunnen worden om de kwaliteit van het onderwijs te verbeteren.

Schoolorganisatiekenmerken (C):

Scholen die meer gebruik hebben gemaakt van ZEBO waren scholen waarin leerkrachten en schoolleiders een meer positieve innovatie-attitude hadden. Onderzocht is ook welke aspecten van de innovatie-attitude een belangrijke rol spelen bij het gebruik van ZEBO. Drie items uit de innovatie-attitudeschaal hebben het gebruik van ZEBO beïnvloed:

- § Het schoolteam gelooft dat het gebruik van ZEBO tot kwaliteitsverbetering zal leiden;
- § Het schoolteam is niet bang voor de veranderingen waartoe ZEBO kan leiden;
- § Leerkrachten denken dat ze invloed kunnen uitoefenen op de maatregelen die genomen worden op basis van ZEBO.

Een ander schoolorganisatiekenmerk dat het gebruik van ZEBO consistent heeft beïnvloed is de innovatiecapaciteit van de school. Scholen met een sterkere innovatiecapaciteit maken over het algemeen meer gebruik van de ZEBO-resultaten dan scholen met een minder ontwikkelde innovatiecapaciteit. Er is onderzocht welke aspecten van de innovatiecapaciteit een rol spelen bij het gebruik van ZEBO. Zes items uit de innovatiecapaciteitschaal hebben het gebruik van ZEBO beïnvloed:

- § De school houdt zicht op de kwaliteit van het functioneren;
- § De school probeert regelmatig uit hoe het onderwijs beter kan;
- § De schoolleider stimuleert de professionele ontwikkeling van het team;
- § Leerkrachten wisselen informatie uit over hun functioneren;
- § Met het hele team is besloten om mee te doen aan het ZEBO-project;
- § De collegialiteit op school is niet groot. Minder collegialiteit in dit onderzoek leidde tot meer ZEBO-gebruik, maar alleen in 2006. De reden hiervoor is niet geheel duidelijk. Het is mogelijk dat ZEBO in deze scholen gebruikt werd om de collegialiteit te verhogen.

Daarnaast werd ZEBO meer gebruikt in scholen die extra tijd en middelen beschikbaar hadden gesteld voor het gebruik van ZEBO. De interviewresultaten laten bijvoorbeeld zien dat vooral leerkrachten soms extra tijd nodig hadden voor het invullen van de ZEBO-vragenlijsten (en voor het organiseren van het invullen van de leerlingvragenlijsten achter de computer). Deze variabele speelde een rol in 2004 en bleef belangrijk in 2006. Het kost tijd en geld om ZEBO te gebruiken, bijvoorbeeld om de resultaten uitgebreid te bespreken, om bepaalde cursussen te volgen, mocht dit nodig zijn op basis van de resultaten, en om nieuwe onderwijsmethoden te implementeren.

Een ander schoolorganisatiekenmerk dat het gebruik van ZEBO beïnvloed heeft is de ZEBO-score. Als een school op of boven het gemiddelde scoort op de procesindicatoren gemeten door ZEBO, dan is minder noodzakelijk om de ZEBO-resultaten te gebruiken, omdat dit inhoudt dat de school het even goed of beter doet dan de gemiddelde Nederlandse school op deze indicatoren. Om hier rekening mee te houden is de ZEBO-score berekend aan de hand van het aantal schalen waarop de school op of boven het

gemiddelde scoorde. De ZEBO-score beïnvloedde alleen het ZEBO-gebruik in 2003 en niet meer in 2004 of 2006. In 2004 en 2006 scoorden veel scholen op of boven het gemiddelde op de ZEBO-schalen, waardoor de variantie tussen scholen minder werd.

Tot slot laten de interviewresultaten zien dat twee andere factoren mogelijk ook het gebruik van ZEBO beïnvloeden:

- § De mate waarin de ZEBO-resultaten overeenkomen met de verwachtingen van de schoolleider. Als de resultaten niet overeenkomen met de verwachtingen van de schoolleider, dan is het mogelijk dat de schoolleider de resultaten afwijst. Dit is dan het einde van de zelfevaluatie;
- § De mate waarin er een open klimaat heerst op de school. Het is belangrijk dat het schoolteam in staat is kritiek te ontvangen. Problemen moeten bespreekbaar zijn. De twee geïnterviewde schoolleiders in 2006 suggereerden dat deze variabele mogelijk een rol speelt bij het gebruik van zelfevaluatie-resultaten. Het is waarschijnlijk dat op scholen waarin men niet in staat is om kritiek te ontvangen en waar problemen niet bespreekbaar zijn, het gebruik van zelfevaluatie-resultaten zeer moeilijk, zo niet onmogelijk wordt.

Conclusie en discussie

In het laatste hoofdstuk van het onderzoek wordt allereerst een reflectie op het onderzoeksdesign gepresenteerd. Een beperking van het onderzoek is dat de steekproef niet representatief is voor denominatie en schoolgrootte, wat de generaliseerbaarheid van de gegevens beperkt.

Ten tweede wordt een aantal praktische implicaties van het onderzoek beschreven. De resultaten van het onderzoek leiden tot verschillende praktische implicaties voor de succesvolle implementatie van SPFS-en in scholen. Het belang van verschillende factoren komt daarbij naar voren, zoals: de betrokkenheid van het hele team bij het besluitvormingsproces m.b.t. de keuze van een aan te schaffen instrument, de duidelijkheid van het doel van de zelfevaluatie voor iedereen, en het belang van stimulatie door de schoolleider om het zelfevaluatieinstrument te gebruiken.

Ten derde worden praktische implicaties voor ontwikkelaars van zelfevaluatie-instrumenten besproken. Zo wordt bijvoorbeeld gesuggereerd dat het gebruik van zelfevaluatie-instrumenten mogelijk zal toenemen wanneer daarbij richtlijnen bij geleverd worden m.b.t. het gebruik van de resultaten van het instrument, inclusief voorbeelden van maatregelen die een school kan nemen om de kwaliteit van het onderwijs te verbeteren.

Ten vierde hebben de resultaten van het onderzoek implicaties voor de theorieontwikkeling over school prestatie feedback. Het SPFS-raamwerk is zeer geschikt gebleken voor de evaluatie van het gebruik van een zelfevaluatieinstrument als

ZEBO. Het was echter niet mogelijk om alle variantie in het gebruik van ZEBO te verklaren met behulp van de variabelen uit het raamwerk. Dit is mogelijk veroorzaakt door de kwaliteit van de instrumenten die gebruikt zijn, maar het is ook mogelijk dat nog een aantal variabelen ontbreekt in het raamwerk. Er worden daarom suggesties gegeven voor variabelen die mogelijk toegevoegd kunnen worden aan het raamwerk, zoals de mate waarin de feedback aanwijzingen bevat over hoe het functioneren van de school verbeteren kan worden, en de mate waarin andere informatie dan ZEBO-informatie de aandacht van de gebruikers opeist. Tot slot wordt gesuggereerd dat persoonlijke kenmerken van de gebruikers (zoals motivatie, locus of control en onzekerheid) mogelijk het gebruik van feedback beïnvloeden.

Het hoofdstuk eindigt met aanbevelingen voor vervolgonderzoek. Het onderzoek heeft bijgedragen aan de kennis over en inzichten in zelfevaluatie, de effecten hiervan en de bevorderende en belemmerende factoren bij zelfevaluatie. De resultaten leidden echter ook, zoals veel onderzoek, tot nieuwe vragen. Een voorbeeld van zo'n vraag is wat de rol van andere data is bij schoolverbeteringsprocessen. Naast zelfevaluatieresultaten hebben scholen andere gegevens om hun functioneren te beoordelen en verbeteren, zoals bijvoorbeeld toets resultaten.

Vragen als: "Hoe kunnen scholen geholpen worden bij het verbeteren en uitbreiden van effectief datagebruik?" vragen om nieuw onderzoek. Verder is nog veel onbekend over het gebruik van school prestatie feedback en de factoren die het gebruik hiervan belemmeren of bevorderen. Zelfevaluatie krijgt over de hele wereld veel aandacht en wereldwijd maken duizenden scholen gebruik van zelfevaluatieinstrumenten. Meer onderzoek naar het gebruik, de effecten en de belemmerende en bevorderende factoren van zelfevaluatie is daarom dringend gewenst.

Appendices

Appendix 3.1 Data collection timeframe

Table 1 *The data collection timeframe*

Time	Data collected
June 2002	Spelling and mathematics tests end grade 3/4 (pre-test)
June 2002	First ZEBO administration (49 schools)
January 2003	First ZEBO administration (15 schools)
January 2003	Spelling and mathematics tests middle grade 4/5 (post-test 1)
May 2003	Evaluation of ZEBO questionnaire
June 2003	Spelling and mathematics tests end grade 4/5 (post-test 2)
December 2003	Interviews with teachers and principals
January - September 2004	Second ZEBO administration (58 schools)
January 2004	Spelling and mathematics tests middle grade 5/6 (post-test 3)
June 2004	Spelling and mathematics tests end grade 5/6 (post-test 4)
May - December 2004	Evaluation of ZEBO questionnaire
January 2005	Spelling and mathematics tests middle grade 6/7 (post-test 5)
May 2005	Interviews with teachers and principals
June 2005	Spelling and mathematics tests end grade 6/7 (post-test 6)
October 2005 - June 2006	Third ZEBO administration (46 schools)
January 2006	Spelling and mathematics tests middle grade 7/8 (post-test 7)
January - August 2006	Evaluation of ZEBO questionnaire
June 2006	Spelling and mathematics tests end grade 7 (post-test 8)
November 2006	Focus group

Appendix 3.2 Evaluation of ZEBO questionnaire

Principal and Teacher versions

The goal of this questionnaire is to evaluate ZEBO use; to explore how ZEBO is used within the school and the effects of that use on education. The following topics are explored: characteristics of ZEBO (A), implementation process features (B), school organisational characteristics (C), ZEBO use (D) and the effects of ZEBO use (E). Confidentiality is assured.
 Please provide the information requested under GENERAL and then complete the questionnaire.
 Most questions / statements are multiple choice. Please choose **one answer**. Some questions are open-ended, please write your answers on the dotted lines provided.
 Please place the completed questionnaire in the enclosed envelop and return.
 Thank you.

GENERAL

Date:

School:

Name:

Function:

Grade:

Date of ZEBO use

A CHARACTERISTICS OF ZEBO

How strongly do you agree with the following statements?

		strongly agree	agree	disagree	strongly disagree	I don't know
A1	The ZEBO output is relevant	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A2	The ZEBO output is up-to-date	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A3	The ZEBO output is inaccurate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A4	ZEBO fits with the needs of our school	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A5	It is easy to enter data in ZEBO	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A6	It is difficult to generate the ZEBO output	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A7	It is easy to alter data in ZEBO	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A8	The ZEBO output is not clear	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A9	Working with ZEBO takes a lot of time	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
A10	ZEBO is easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B IMPLEMENTATION PROCESS FEATURES

B1a How many hours of ZEBO implementation training and support (internal and external) did you receive, e.g. from the school advisory service?

none

IF YOU CHOSE "NONE" PLEASE PROCEED TO QUESTION C5

- 1-4 hours
- 5-10 hours
- 11-20 hours
- 21-30 hours
- >30 hours

How strongly do you agree with the following statements?

		<i>strongly agree</i>	<i>agree</i>	<i>disagree</i>	<i>strongly disagree</i>	<i>I don't know</i>
B1b	Our school received sufficient training in the implementation of ZEB0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B2	I am not satisfied with the content of the training	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B3	Our school received sufficient support in the implementation of ZEB0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B4	I am not satisfied with the content of the support	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B5	In the principal questionnaire: I encouraged ZEB0 use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B5	In the teacher questionnaire: The principal encouraged ZEB0 use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B6	We felt pressured to implement ZEB0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B7	The goal of the implementation of ZEB0 is not clear	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C SCHOOL ORGANISATIONAL CHARACTERISTICS

How strongly do you agree with the following statements?

		<i>strongly agree</i>	<i>agree</i>	<i>disagree</i>	<i>strongly disagree</i>	<i>I don't know</i>	<i>does not apply here</i>
C1a	I think that ZEB0 use will improve the quality of our school	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C1b	I think that ZEB0 use will have a negative influence on my work	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C1c	I am afraid a lot of things will change because of ZEB0 use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C2	Our school reserved extra time and resources for the use of ZEB0	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3a	As a team we decided to participate in the ZEB0-project	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3b	In the principal questionnaire: Teachers can influence the measures taken as a result of the ZEB0 output, to improve the quality of education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3b	In the teacher questionnaire: I can influence the measures taken as a result of the ZEB0 output, to improve the quality of education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3c	Our school monitors the quality of education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3d	In the principal questionnaire: I take the wishes and needs of the teachers into account	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3d	In the teacher questionnaire: The principal takes the wishes and needs of the teachers into account	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3e	In the principal questionnaire: I encourage the professional development of teachers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3e	In the teacher questionnaire: The principal encourages the professional development of teachers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

		strongly agree	agree	disagree	strongly disagree	I don't know	does not apply here
Our school							
C3f	is able to improve its quality independently	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3g	experiments regularly with how to improve education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3h	does not change, unless it must	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3i	is aimed at a continuous improvement of its functioning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3j	is accustomed to frequent evaluation of its functioning	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
In our school							
C3k	teachers take extra courses even when it is not obligatory	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3l	learning by teachers is very important	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3m	the team cohesion is not strong	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3n	we feel responsible, as a team, for our education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3o	teachers actively work together and across classes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C3p	teachers exchange information on their functioning within the class	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

D ZEBO USE

Did you study the following analyses from ZEBO?

An overview of each separate ZEBO scale for:		yes	no
D1a	the teachers and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>
D1b	the pupils and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>
D1c	the principal and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>
An overview in tables or graphics of how the school is judged by:			
D1d	the teachers in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>
D1e	the pupils in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>
D1f	the principal in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>
D1g	A discrepancy analysis: a textual overview of the results of the school and differences between the opinions of the pupils and teachers and between the teachers and the principal	<input type="checkbox"/>	<input type="checkbox"/>

Were measures taken on the basis of the following ZEBO output:

An overview of each separate ZEBO scale for:		yes	no	I don't know
D2a	the teachers and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2b	pupils and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2c	the principal and differences from the national mean	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
An overview in tables or graphics of how the school is judged by:				
D2d	the teachers in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2e	pupils in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D2f	the principal in comparison with other schools in the Netherlands	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dg	A discrepancy analysis: a textual overview of the results of the school and differences in opinions between the pupils and teachers and between the teachers and the principal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

To what degree do the following statements apply to your school

		To a great degree	To a moderate degree	To a small degree	To a minimal degree	I don't know	does not apply here
D3	The ZEBO output was discussed within the school	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D4a	ZEBO use provided me with new insights	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D4b	if yes, can you give an example of a new insight	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
.....							
.....							
.....							

		To a great degree	To a moderate degree	To a small degree	To a minimal degree	I don't know	does not apply here
D5	The ZEBO output highlighted certain problems within the school	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D6a	School staff devised solutions for the problems highlighted by ZEBO	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D6b	I devised solutions for the problems highlighted by ZEBO	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7a	On the basis of the ZEBO output, school staff took measures to improve the quality of education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7b	If yes, can you give an example						
						
						
						
						
D7c	On the basis of the ZEBO output I took measures to improve the quality of education	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D7d	If yes, can you give an example						
						
						
						
						

E EFFECTS OF ZEBO USE

To what degree do you think that because of ZEBO use:

		To a great degree	To a moderate degree	To a small degree	To a minimal degree	I don't know	does not apply here
E2.1.a	more consultation on the functioning of the school and on the quality of education occurs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.1.b	If yes, can you give an example?						
						
						
						
						
E2.2.a	In the principal questionnaire your functioning as an educational leader improved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.2.b	In the teacher questionnaire the functioning of the principal improved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	If yes, can you give an example?						
						
						
						
						

Appendices

		To a great degree	To a moderate degree	To a small degree	To a minimal degree	I don't know	does not apply here
E2.3.a	there is more attention to your professional development	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.3.b	If yes, can you give an example?						
						
						
						
						
						
E2.4.a	the achievement orientation has been enhanced	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.4.b	If yes, can you give an example?						
						
						
						
						
						
E2.5.a	the team cohesion is stronger	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.5.b	If yes, can you give an example?						
						
						
						
						
						
E2.6.a	pupil care has improved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.6.b	If yes, can you give an example?						
						
						
						
						
						
E2.7.a	your teaching has improved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.7.b	If yes, can you give an example?						
						
						
						
						
E2.8.a	pupil achievement is evaluated on a more regular basis	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.8.b	If yes, can you give an example?						
						

E2.9.a	adaptive education has improved	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E2.9.b	If yes, can you give an example?						
						
						
						
						
						

E3.a	did ZEBO use have any negative effects?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E3.b	If yes, can you give an example?						
						
						
						
						
						

Thank you for completing the questionnaire
Kim Schildkamp

Appendix 3.3 Reliability of the Evaluation of ZEBO questionnaire, 2004 and 2006

Table 1 Reliability of the questionnaire scales at teacher and principal level, 2004 and 2006

Reliability at	2004 PL Cronbach's α (items)	2004 TL Cronbach's α (items)	2004 ASL Cronbach's α (items)	2006 PL Cronbach's α (items)	2006 TL Cronbach's α (items)	2006 ASL Cronbach's α (items)
Scale	(N=48)	(N=236)	(N=50)	(N=25)	(N=141)	(N=31)
Characteristics of ZEBO	0.73 (9)	0.80 (9)	0.81 (9)	0.80 (9)	0.64 (9)	0.86 (9)
Implementation process features: training and support	Too many missing cases	0.89 (4)	0.62 (4)	0.63 (3)	0.72 (5)	0.86 (4)
Implementation process features: Pressure and promoting factors	0.65 (3)	0.59 (3)	0.75 (3)	0.75 (3)	0.50 (2)	0.80 (3)
School organisational features: innovation attitude	0.42 (7)	0.68 (7)	0.62 (7)	0.60 (7)	0.76 (7)	0.85 (7)
School organisational features: innovation capacity	0.78 (12)	0.84 (12)	0.89 (12)	0.79 (12)	0.61 (10)	0.78 (10)
ZEBO use	0.81 (9)	0.83 (9)	0.81 (9)	0.85 (9)	0.82 (9)	0.88 (9)
Conceptual use of ZEBO	0.72 (4)	0.75 (4)	0.75 (5)	0.83 (4)	0.68 (4)	0.79 (4)
Instrumental use of ZEBO	0.66 (5)	0.75 (5)	0.67 (5)	0.68 (5)	0.64 (4)	0.70 (5)
Effects of the use of ZEBO	0.90 (9)	0.89 (9)	0.88 (9)	0.92 (9)	0.95 (9)	0.76 (9)

Notes: PL: principal level, TL: teacher level, ASL: aggregated school level.

Appendix 3.4 Questionnaire response rates

Table 1 *Functions of the returning respondents and the numbers of questionnaires returned*

Function	2003 Number of questionnaires (%)	2004 Number of questionnaires (%)	2006 Number of questionnaires (%)
Principal	42 (16)	48 (17)	24 (14)
Assistant principal	3 (1)	2 (1)	1 (1)
Teacher grade 1	14 (5)	7 (2)	3 (2)
Teacher grade 2	9 (3)	13 (5)	4 (2)
Teacher grade 3	32 (12)	20 (7)	18 (11)
Teacher grade 4	26 (10)	24 (8)	12 (7)
Teacher grade 5	11 (4)	21 (7)	10 (6)
Teacher grade 6	18 (7)	19 (7)	10 (6)
Teacher grade 7	12 (5)	24 (8)	12 (7)
Teacher grade 8	18 (7)	23 (8)	17 (10)
Teacher grade 0/1/2	20 (7)	31 (11)	18 (11)
Teacher grade 1/2/3/4	2 (1)	1 (0)	0 (0)
Teacher grade 2/3	0 (0)	1 (0)	0 (0)
Teacher grade 3/4	8 (3)	5 (2)	11 (7)
Teacher grade 3/4/5	1 (0)	4 (1)	0 (0)
Teacher grade 4/5	7 (3)	1 (0)	2 (1)
Teacher grade 5/6	12 (5)	13 (5)	8 (5)
Teacher grade 6/7	4 (2)	5 (2)	4 (2)
Teacher grade 6/7/8	2 (1)	1 (0)	0 (0)
Teacher grade 7/8	10 (4)	6 (2)	8 (5)
Internal educational advisor & teacher	3 (1)	11 (4)	4 (2)
Substitute teachers (grade 1 to 8)	4 (2)	2 (1)	0 (0)
ICT assistant and teacher	2 (1)	2 (1)	0 (0)
Interim principal	1 (1)	0 (0)	0 (0)
Total	261	284	166

Note: (%) refers to percentage of questionnaires which were returned.

Table 2 *Number of questionnaires returned per school per year*

Number of questionnaires returned	Number of schools (2003)	Number of schools (2004)	Number of schools (2006)
1	4	3	4
2	5	4	1
3	3	3	1
4	7	5	3
5	5	8	8
6	10	7	2
7	7	6	5
8	5	6	5
9	4	8	2
Total	50	50	31

Appendix 3.5 Interview Schedule

Interview "Evaluation of the Use of ZEBO

Respondent Name:

Date of interview:

Introduction

Before we start, may I have your permission to tape this interview?

Confidentiality is assured. Recently you completed a questionnaire on the use of ZEBO. The goal of this interview is to gain more information about the use of ZEBO within your school. The interview will take 30 minutes to one hour. Before we start, do you have any questions?

The first couple of questions deal with the characteristics of ZEBO and how you judge these characteristics.

1. a. Are there any characteristics of ZEBO that you judge positively?
b. If yes, which?
c. Why?
2. a. Are there any characteristics of ZEBO that you judge negatively?
b. If yes, which?
c. Why?
3. a. Which information from ZEBO do you value most?
b. Why?

The next questions deal with the implementation of ZEBO

4. With what goal was ZEBO implemented?
5. Were you involved in the decision to participate in the ZEBO-project?
6. a. Did specific circumstances promote or obstruct the implementation of ZEBO?
b. If yes, which circumstances?
c. How?
7. a. Did the principal (or when the respondent is a principal: did you) encourage the use of ZEBO?
b. If yes, in which way?
8. a. Did you receive any training and support in the implementation of ZEBO?
b. If yes: was the training and support sufficient?
c. If no: did you feel the need for training and support?

The next questions deal with characteristics of the school organisation

9. a. Does the school carry out activities to measure and improve the quality of education (besides ZEBO)?
b. If yes, can you give some examples?
10. Did the school make extra facilities, like money and personnel, available for the implementation of ZEBO?
11. a. Before using ZEBO, did you have a positive attitude to the implementation of ZEBO?
b. And after having used it, do you have a positive attitude to ZEBO?

This question was added in the interview schedule for 2005

12. a. How does the school work on the team's professional development?
- b. What is the role of the principal in this?

The next questions deal with the use of ZEBO

13. a. Did certain problems arise during the use of ZEBO?
- b. If yes, which problems?
14. a. Which ZEBO output did you see?
- b. Did you find these results clear?

This question was added in the interviews for 2005

15. a. Did the school use the ZEBO output the second time differently to the first time?
- b. If yes, how and why?

This question was added in the interviews for 2005

16. a. Did the school use the ZEBO output the second time less, more, or the same as the first time?
- b. If more or less, how and why?

This question was added in the interviews for 2005

17. a. Did you compare the output of the second time to the output of the first time?
- b. If yes, did the output differ?
- c. If yes, what do you think caused these differences?
- d. What was the goal of comparing the outputs?
18. a. Was the output discussed?
- b. If yes, how many times?
- c. With whom?
- d. What was discussed?
19. a. Did the ZEBO output highlight certain problems within the school?
- b. Were these problems familiar?
20. a. Did the use of ZEBO provide you with new insights?
- b. If yes, what were they?
21. a. Were measures taken on the basis of the ZEBO output?
- b. If yes, what were they?
- c. With what goal?

That was the last question. Thank you very much for your cooperation.

Appendix 3.6 Topic list for the focus group

Welcome to this ZEBO meeting and thank you for taking the time to participate. As I told you before, I would like to discuss some of my conclusions with you. I invited you here because you have all used ZEBO over the last five years. I would like to commence with everybody introducing him or herself and telling us about your experiences with ZEBO.

- § The results of this study show that some schools make use of the ZEBO output to improve the quality of education, but most schools are not using the ZEBO output (effectively). Self-evaluation appears for most schools to be a difficult task. Do you have any idea as to why most schools appear to be unable to use self-evaluation results to make improvements?
- § Did you come across certain obstacles in using ZEBO and the ZEBO results?
- § Do you think that it is due to certain characteristics of ZEBO that most schools are not using the ZEBO output?
- § Do you think that schools need training and support in the use of the ZEBO results? If yes, what kind of training and support?
- § The results of my study show that the principal plays an important role in the use of the ZEBO output. Is it possible that principals in some schools inhibit the use of the ZEBO output? If yes, why and how?
- § Teachers also play an important role in the use of ZEBO. Is it possible that teachers in some schools inhibit the use of the ZEBO output? If yes, why and how?

Next, I would like to put forward some statements. Can you express, for each of these statements, whether you agree or disagree and why.

1. We use ZEBO mostly for the Inspectorate.
2. The feedback from ZEBO is more useful than the inspectorate information.
3. Schools which score on or above average on the ZEBO use scales do not have to use the ZEBO output.
4. It is not easy to find solutions for the problems which ZEBO highlighted.
5. We do not have enough time to use the ZEBO output.
6. The team is able to use the ZEBO output to make improvement without the help and support of the principal.
7. We need extra time and resources to make use of the ZEBO output.
8. Teachers think that the use of ZEBO is competing with their other tasks.
9. Teachers with more experience perceive ZEBO more negatively.
10. We need training and support in the use of the ZEBO output.
11. In England, several schools are receiving the help and support of a “critical friend” in their self-evaluation processes. A “critical friend” is a person from outside the school who supports a school in the self-evaluation process but also provides the school with a critical view. We could also benefit from the help of a “critical friend”.

Is there anything we have not discussed this afternoon, but which you would like to mention or add to our discussion?

Thank you very much for your participation. I will send a copy of my notes to your school for approval.

Appendix 3.7 Functions of the respondents interviewed

Table 1 *Functions of the respondents interviewed in 2003 across LoSE, AvSE, and HiSE schools*

Use	Teacher grade 3	Teacher grade 4	Teacher grade 3/4	Teacher grade 4/5	EA ¹	AP ² /5/6	AP/3/IEA	Principal	Total
LoSE	1	1 ³						1	3
LoSE			1		1			1	3
LoSE	1		1					1	3
LoSE	1			1				1	3
AvSE			1 ³			1		1	3
AvSE	1	1						1	3
AvSE	1	1						1	3
AvSE	1	1						1	3
AvSE	1	1						1	3
HiSE		1					1		2
HiSE	1 ³							1	2
Total	7	5	2	1	1	1	1	10	31

IEA: Internal Educational Advisor; AP: Assistant Principal.

Table 2 *Functions of the respondents interviewed in 2005*

Use	Teacher grade 5	Teacher grade 6	Teacher grade 5/6	Teacher grade 7	AP/ Teacher grade 4/6	IEA/ teacher	AP/IEA/ teacher	Principal	Total
LoSE	1	1					1	1	4
LoSE	1			1	1				3
LoSE		1						1	2
AvSE	1	1							2
AvSE			1					1	2
AvSE			1			1		1	3
HiSE	1	1						1	3
HiSE	1			1				1	3
HiSE	1 ³	1						1	3
Total	6	4	2	2	1	1	1	7	25

IEA: Internal Educational Advisor; AP: Assistant Principal.

¹ Internal Educational Advisor.

² Assistant Principal.

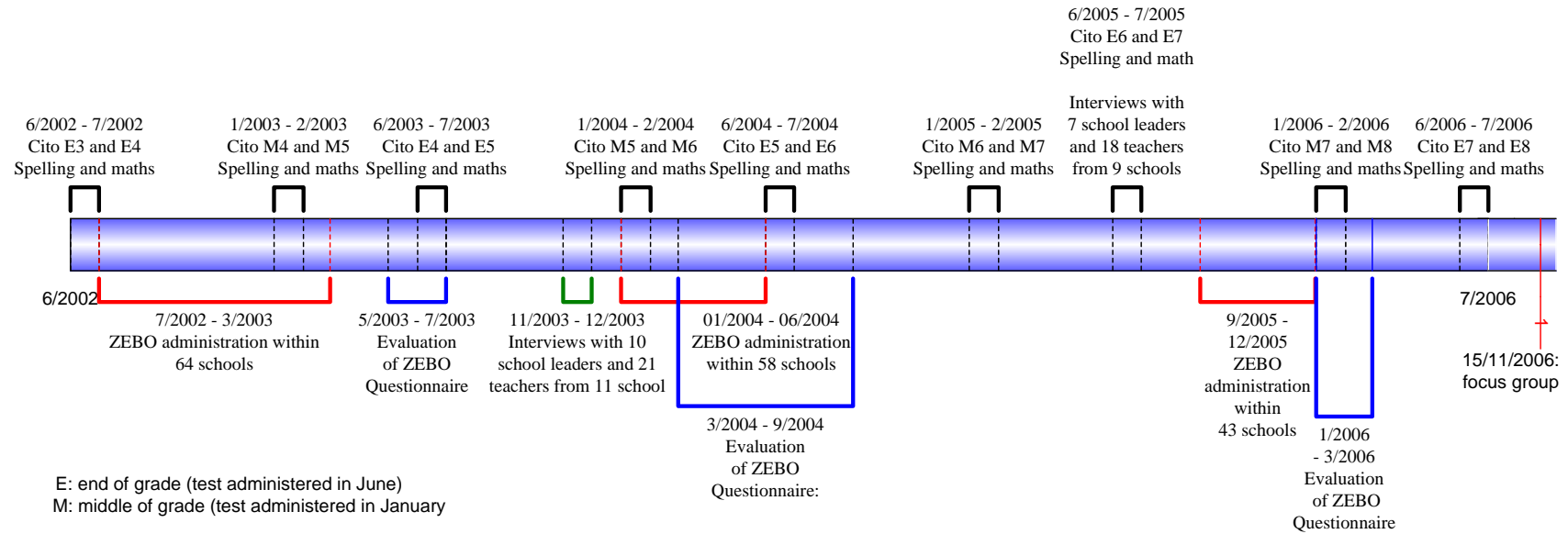
³ Internal Educational Advisor.

Appendix 3.8 Matrix of the factors studied, the instruments used and administration dates

Table 1 *The factors studied and the instruments used and administration dates*

Factors	Principal and teacher questionnaires	Achievement tests	Interviews with school staff	ZEBO output
<i>Characteristics of ZEBO</i>	2003		2003	
	2004		2004	
	2006		2006	
<i>Implementation process features</i>	2003		2003	
	2004		2004	
	2006		2006	
<i>School organisational characteristics</i>	2003			2003
	2004			2004
	2006			2006
<i>ZEBO use</i>	2003			
	2004			
	2006			
<i>Effects of the use of ZEBO on pupil achievement</i>		2001-2006		
<i>Other effects of the use of ZEBO</i>	2003			
	2004			
	2006			

Appendix 3.9 Timeline of the study



Appendix 3.10 Interview codes

Table 1 *Families and number of codes used*

Families and categories	Number of codes
<i>Characteristics of ZEBO as perceived by its users</i>	
Positive characteristics of ZEBO	14
Negative characteristics of ZEBO	13
Most valuable information/output from ZEBO	7
<i>Implementation process features</i>	
Training and support	11
Decision to participate in the ZEBO project	18
Decision on what happens with the results	4
Problems with ZEBO	6
The role of the principal	11
Goal of using ZEBO	11
Goal of using ZEBO for the second time	2
Circumstances surrounding the implementation of ZEBO	10
<i>School organisational characteristics</i>	
Time and resources available for ZEBO implementation	4
Quality care activities	10
Professional development of school staff	4
Professional development and the principal	3
<i>ZEBO use</i>	
Results were studied	3
Clarity of results	3
Results of the first and second use were compared	2
New insights	9
The results highlighted problems	7
Results were discussed	26
Measures taken on the basis of the results ZEBO output	25
Differences from the first use	7
Total number of codes	176

Appendix 3.11 Inter-rater agreement

Table 1 *Inter-rater agreement on coding*

Category	Cohen's Kappa
Positive characteristics of ZEBO	0.93
Negative characteristics of ZEBO	0.71
Most valuable information from ZEBO	0.89
Goal of ZEBO	0.90
Involvement in decision to participate	0.83
Circumstances surrounding the implementation	1.00
Encouragement by the principal	0.91
Training and support	1.00
Quality care activities	1.00
Extra facilities for ZEBO	1.00
Problems with the use of ZEBO	1.00
Studied the results	0.66
Discussed the results	0.82
The results highlighted problems	1.00
The results led to new insights	0.79
Measures were taken based on the results	0.89

Appendix 4.1 Comparison of the results from the interviews with “LoSE”, “AvSE”, and “HiSE” schools

The results from the interviews on the use of ZEB0 may be found in Table 1 (2003) and Table 2 (2005). Behind the different answers are two numbers in brackets. The first number represents the number of respondents who gave that answer. The second number represents the number of schools from which these respondents came.

Table 1 *The use of ZEB0 in LoSE, AvSE and HiSE schools (2003)*

Respondents from LoSE schools (12 respondents: 4 schools)	Respondents from AvSE schools (15 respondents: 5 schools)	Respondents from HiSE schools (4 respondents: 2 schools)
<i>Did you see the results from ZEB0?</i>	<i>Did you see the results from ZEB0?</i>	<i>Did you see the results from ZEB0?</i>
§ Yes (9:3)	§ Yes (15:5)	§ Yes (4:2)
§ No (3:1)		
<i>Were the results clear?</i>	<i>Were the results clear?</i>	<i>Were the results clear?</i>
§ Yes: (8:2)	§ Yes (14:5)	§ Yes (4:2)
§ I would have liked more explanation (4:2)	§ I would have liked more explanation (1:1)	
<i>Were the results discussed?</i>	<i>Were the results discussed?</i>	<i>Were the results discussed?</i>
§ No (6:2)	§ No (1:1)	§ The results were discussed in the participation council (1:1)
§ In the team meeting the best and worst scores were discussed (2:1)	§ Yes, the results from the individual teachers were discussed in performance interviews (1:1)	§ In the team meeting the best and worst scores were discussed (2:1)
§ The results were discussed between the principal and the individual teachers (3:1)	§ The results were discussed point by point (2:1)	§ The results were discussed point by point (3:2)
§ The results were discussed in a team meeting (4:2)	§ The results were discussed in several team meetings (13:5)	§ The results were discussed in several team meetings (4:2)
	§ The differences with the national mean were discussed (6:2)	§ We discussed how we should deal with the school management
	§ Differences in opinion were discussed (6:2)	
	§ The positive scores were discussed (1:1)	
<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>
§ Some scores were below average (2:2)	§ Some scores were below average (7:3)	§ Some scores were below average (2:1)
§		
<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>
		§ The school management team was judged very negatively by the teachers (2:1)
<i>Did the results lead to new insights?</i>	<i>Did the results lead to new insights?</i>	<i>Did the results lead to new insights?</i>
§ No (11:4)	§ No (12:5)	§ No (3:2)
§ Yes, some pupils find my lessons boring (1:1)	§ How we score in comparison to the national mean (2:1)	§ How we score in comparison to the national mean (1:1)
	§ Our quality care is better than expected (1:1)	

Table 1 *The use of ZEB0 in LoSE, AvSE and HiSE schools (2003) continued*

Respondents from LoSE schools (12 respondents: 4 schools)	Respondents from AvSE schools (15 respondents: 5 schools)	Respondents from HiSE schools (4 respondents: 2 schools)
<i>Were certain measures taken based on the results?</i>	<i>Were certain measures taken based on the results?</i>	<i>Were certain measures taken based on the results?</i>
§ No (11:4)	§ No (6:3)	§ We are consulting with the school management team how to improve the relationship with the teachers and the school management team (2:1)
§ I developed an action plan (1:1)	§ I developed an action plan (2:2)	§ I compared the results from ZEB0 with the results from a quality care instrument that we used last year (1:1)
	§ We started with classroom consultation (3:1)	§ We started with classroom consultation (1:1)
	§ I spend more time on the low achieving pupils (1:1)	
	§ Participation in a project to stimulate adaptive education (3:1)	
	§ I used the ZEB0 output for the school plan (1:1)	
	§ I developed a policy to decrease the workload of the teachers (1:1)	
	§ We are trying to develop a shared vision (1:1)	
	§ I am working on a policy to improve quality care (1:1)	
	§ Improving independent learning by means of block teaching (1:1)	

Table 2 *The use of ZEB0 in LoSE, AvSE and HiSE schools (2005)*

Respondents from LoSE schools (4 respondents: 1 school)	Respondents from AvSE schools (12 respondents: 5 schools)	Respondents from HiSE schools (9 respondents: 3 schools)
<i>Did you see the results from ZEB0?</i>	<i>Did you see the results from ZEB0?</i>	<i>Did you see the results from ZEB0?</i>
§ Yes (4:1)	§ Yes (12:5)	§ Yes (9:3)
<i>Were the results clear?</i>	<i>Were the results clear?</i>	<i>Were the results clear?</i>
§ Yes: (1:1)	§ Yes (12:5)	§ Yes (9:3)
§ Complicated (1:1)		
§ Did not concur with our expectations (2:1)		
<i>Were the results discussed?</i>	<i>Were the results discussed?</i>	<i>Were the results discussed?</i>
§ No (1:1)	§ No (1:1)	§ The results were discussed in the quality care project group (3:1)
§ The results were discussed in (several) team meeting (3:1)	§ The results were discussed in (several) team meetings (7:3)	§ The results were discussed in (several) team meetings (9:3)
§ Differences within the team were discussed (1:1)	§ Differences within the team were discussed (1:1)	§ Differences within the team were discussed (4:3)
§ The negative results from the principal were discussed (2:1)	§ The results were discussed in performance interviews (2:1)	§ The results were discussed in performance interviews (1:1)
	§ The differences with the national mean were discussed (6:4)	§ The differences with the national mean were discussed (6:3)
	§ The remarkable or extraordinary items were discussed (6:3)	§ The remarkable or extraordinary items were discussed (3:2)
	§ The items that ask for improvement were discussed (5:4)	§ We discussed how we may improve on certain items (2:2)
	§ In project groups we worked on different items from ZEB0 (1:1)	§ The internal educational advisor discussed the results in terms of what the results mean for our school (1:1)
	§ Do we agree or disagree with the results (1:1)	§ Teachers expressed their opinion about the results (1:1)
	§ The principal discussed the classroom results individually with the teachers (2:2)	§ The results were discussed with the participation council (1:1)
	§ The results were discussed with the school advisory service (2:1)	§ In the team meeting the good and poor scores were discussed (4:2)
		§ The positive items were discussed (3:3)
		§ In the parents' council (1:1)

Table 2 *The use of ZEB0 in LoSE, AvSE and HiSE schools (2005) continued*

Respondents from LoSE schools (4 respondents: 1 school)	Respondents from AvSE schools (12 respondents: 5 schools)	Respondents from HiSE schools (9 respondents: 3 schools)
<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>	<i>Did the results point out certain problems within the school?</i>
§ Some scores were below average (2:1)	§ Some scores were below average (10:5)	§ Some scores were below average (8:3)
§ No (1:1)	§ No (1:1)	§ I do not know (1:1)
§ The principal was judged very negatively (3:1)	§ I do not know (1:1)	
<i>Did the results lead to new insights?</i>	<i>Did the results lead to new insights?</i>	<i>Did the results lead to new insights?</i>
§ No (2:1)	§ No (8:5)	§ No (4:2)
§ I do not know (1:1)	§ Differences with the national mean (2:2)	§ Differences within the team (3:2)
§ Yes, the negative results from the principal (1:1)	§ On some items we scored very positively (1:1)	§ There were some communication problems (1:1)
	§ How pupils perceive my lessons (1:1)	§ On some items we scored very positively (1:1)
<i>Were certain measures taken based on the results?</i>	<i>Were certain measures taken based on the results?</i>	<i>Were certain measures taken based on the results?</i>
§ No (1:1)	§ No (3:2)	§ The principal communicates her activities more (2:1)
§ We have started working on classroom management (1:1)	§ We have developed an action list (3:1)	§ We have developed an action list (1:1)
§ We have made changes in the way we organize team meetings (1:1)	§ School staff has participated in professional development courses (2:1)	§ School staff participated in professional development courses (2:1)
	§ Translate items in desired teacher behaviour (1:1)	§ We agreed to use the results for school improvement (1:1)
	§ Establish, retain and evaluate our quality regularly (1:1)	§ We have started a quality care project group (3:1)
	§ We have been working on clarifying our school's objectives (1:1)	§ We have made a protocol for more concordance within the team (1:1)
	§ Evaluation has become a central issue (1:1)	§ Re-dividing teachers' workload (1:1)
	§ More cooperation between staff (1:1)	§ The results have been used in making personal development plans (3:1)
<i>The second ZEB0 use?</i>	<i>The second ZEB0 use?</i>	<i>The second ZEB0 use?</i>
§ I do not know whether or not the second ZEB0 use differed from the first time (2:1)	§ I do not know whether or not the second ZEB0 use differed from the first time (1:1)	§ We have made more use of the results the second time (3:2)
§ The second ZEB0 use was the same as the first time (2:1)	§ The second ZEB0 use was the same as the first time (1:1)	§ The second ZEB0 use was the same as the first time (5:3)
§ The second ZEB0 we took no measures based on the results because we interpreted some questions wrongly (1:1)	§ The second time we have started working with the results that were remarkable or extraordinary (1:1)	§ It was easier to use the second time, because we knew what to expect (3:3)
§ The results of the first and second ZEB0 use were not compared (2:1)	§ We have asked the school advisory service to help us in using the results (3:1)	§ It did not feel as a one time only experience anymore (1:1)
§ We compared the results of the first and second ZEB0 use to see whether or not differences existed (2:1)	§ We have compared the results of the first and second ZEB0 use to see whether or not differences existed (2:1)	§ We have compared the results of the first and second ZEB0 use to see whether or not differences existed (9:3)
	§ We have worked on other items than the first time (1:1)	§ The entire school has used ZEB0 the second time (1:1)
	§ The second time we took less measures based on the ZEB0-results because of impeding circumstances, but the results did not indicate any problems (2:1)	§ More discussion has taken place based on the second results (1:1)
	§ The results of the first and second ZEB0 use were not compared (7:3)	
	§ The first ZEB0 use gave us a rough expression. The second time we have extracted certain items on which we want to improve (2:2)	

Appendix 5.1 Multilevel analyses of the effects of conceptual and instrumental ZEB0 use on pupil achievement on spelling in grade 4 (2003)

Table 1 *Effects of conceptual and instrumental use of ZEB0 on the spelling test in Grade 4*

	Model CON0		Model CON1		Model INSTR0		Model INSTR1	
	<i>(N=722)</i>		<i>(N=741)</i>		<i>(N=741)</i>		<i>(N=741)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Intercept	-0.28	0.23	-0.27	0.23	-0.27	0.23	-0.40	0.40
Girls vs. Boys	0.21	0.06	0.21	0.06	0.21	0.06	0.21	0.06
Combination class: yes vs. no	-0.27	0.11	-0.28	0.11	-0.28	0.11	-0.28	0.11
SES (1.25 vs. 1.0)	-0.07	0.08	-0.07	0.08	-0.07	0.08	-0.07	0.08
SES (1.9 vs. 1.0)	0.04	0.12	0.05	0.12	0.05	0.12	0.05	0.12
Av. vs. low IQ	0.33	0.09	0.33	0.09	0.33	0.09	0.33	0.09
High vs. low IQ	0.89	0.11	0.89	0.11	0.89	0.11	0.89	0.11
Pre-test grade 3	0.42	0.04	0.42	0.03	0.42	0.03	0.42	0.03
Conceptual use in 2003			0.01	0.03				
Instrumental use in 2003							0.00	0.00
Variance components								
Between classes	0.09	0.03	0.09	0.03	0.09	0.03	0.09	0.03
Between pupils	0.55	0.03	0.54	0.03	0.54	0.03	0.54	0.03
Percentage explained								
Between classes			3.93%				3.00%	
Between pupils			0.47%				0.47%	
Deviance	1669		1701		1701		1701	
Improvement in model fit (p)			1.00				0.77	

Appendix 5.2 Multilevel analyses of the effects of conceptual and instrumental ZEBU use on pupil achievement on mathematics in grade 4 (2003)

Table 1 *Effects of conceptual and instrumental use of ZEBU in 2003 on the mathematic test in Grade 4*

	ModelCON0		ModelCON1		Model INSTR0		Model INSTR1	
	<i>(N=722)</i>		<i>(N=722)</i>		<i>(N=741)</i>		<i>(N=741)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Intercept	-0.25	0.11	0.02	0.38	-0.25	0.10	0.03	0.32
Girls vs. Boys	-0.15	0.04	-0.15	0.04	-0.15	0.04	-0.15	0.04
SES (1.25 vs. 1.0)	0.01	0.06	0.01	0.06	0.01	0.06	0.01	0.06
SES (1.9 vs. 1.0)	-0.23	0.10	-0.23	0.10	-0.23	0.10	-0.23	0.10
Av. vs. low IQ	0.40	0.07	0.40	0.07	0.40	0.07	0.40	0.07
High vs. low IQ	0.86	0.09	0.86	0.09	0.85	0.09	0.85	0.09
Pre-test grade 3	0.59	0.03	0.59	0.03	0.59	0.03	0.59	0.03
Conceptual use in 2003			-0.02	0.03				
Instrumental use in 2003							-0.02	0.02
Variance components								
Between classes	0.10	0.03	0.10	0.03	0.09	0.02	0.09	0.02
Between pupils	0.34	0.02	0.34	0.02	0.34	0.02	0.34	0.02
Percentage explained								
Between classes			0.96%				0.00%	
Between pupils			0.23%				0.00%	
Deviance	1413		1412		1438		1437	
Improvement in model fit (p)			0.47				0.35	

Appendix 5.3 Multilevel analyses of the effects of conceptual and instrumental ZEB0 use on pupil achievement on spelling in grade 5 (2003)

Table 1 *Effects of conceptual and instrumental use of ZEB0 in 2003 on the spelling test in Grade 5*

	Model CON0		ModelCON1		Model INSTR0		Model INSTR1	
	<i>(N=722)</i>		<i>(N=722)</i>		<i>(N=741)</i>		<i>(N=741)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Intercept	-0.65	0.13	-0.73	0.37	-0.65	0.12	-0.26	0.35
Girls vs. Boys	0.22	0.05	0.22	0.05	0.21	0.05	0.21	0.05
Language at home (Dutch vs. Turkish)	0.12	0.13	0.12	0.13	0.11	0.13	0.11	0.13
Language at home (Dutch vs. Dialect)	-0.34	0.19	-0.34	0.19	-0.34	0.19	-0.33	0.19
Language at home (Dutch versus other)	0.00	0.15	0.00	0.15	-0.04	0.15	-0.04	0.15
Av. vs. low IQ	0.36	0.08	0.36	0.08	0.37	0.08	0.36	0.08
High vs. low IQ	0.75	0.10	0.75	0.10	0.78	0.10	0.78	0.10
Pre-test grade 4	0.56	0.03	0.56	0.03	0.55	0.03	0.55	0.03
Conceptual use in 2003			0.01	0.03				
Instrumental use in 2003							-0.02	0.02
Variance components								
Between classes	0.12	0.03	0.12	0.03	0.12	0.03	0.12	0.03
Between pupils	0.47	0.03	0.47	0.03	0.47	0.03	0.47	0.03
Percentage explained								
Between classes			2.30%				0.79%	
Between pupils			0.52%				0.17%	
Deviance	1643		1643		1687		1686	
Improvement in model fit (p)			0.91				0.24	

Appendix 5.4 Multilevel analyses of the effects of the conceptual and instrumental ZEB0 use on pupil achievement on mathematics in grade 5 (2003)

Table 1 *Effects of conceptual and instrumental use of ZEB0 on the mathematics test in Grade 5*

	Model CON0		ModelCON1		Model INSTR0		Model INSTR1	
	<i>(N=722)</i>		<i>(N=722)</i>		<i>(N=741)</i>		<i>(N=741)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.	Est.	S.E.	Est.	S.E.
Intercept	-0.20	0.10	-0.74	0.44	-0.20	0.10	-0.47	0.43
SES (1.25 vs. 1.0)	-0.13	0.06	-0.12	0.06	-0.13	0.06	-0.13	0.06
SES (1.9 vs. 1.0)	-0.27	0.15	-0.26	0.15	-0.26	0.15	-0.26	0.15
Language at home (Dutch vs. Turkish)	-0.14	0.16	-0.13	0.16	-0.14	0.16	-0.14	0.16
Language at home (Dutch vs. Dialect)	-0.19	0.16	-0.19	0.16	-0.19	0.16	-0.19	0.16
Language at home (Dutch versus other)	-0.39	0.17	-0.38	0.17	-0.41	0.16	-0.41	0.16
Av. vs. low IQ	0.29	0.08	0.29	0.08	0.30	0.07	0.30	0.07
High vs. low IQ	0.62	0.10	0.62	0.10	0.62	0.10	0.62	0.10
Pre-test grade 4					0.57	0.03	0.57	0.03
Conceptual use in 2003			0.04	0.03				
Instrumental use in 2003							0.02	0.03
Variance components								
Between classes	0.21	0.05	0.21	0.05	0.21	0.05	0.21	0.05
Between pupils	0.37	0.02	0.37	0.02	0.37	0.02	0.37	0.02
Percentage explained								
Between classes			1.35%				0.92%	
Between pupils			0.51%				0.34%	
Deviance	1440		1438		1476		1476	
Improvement in model fit (p)			0.21				0.52	

Appendix 5.5 Multilevel analyses with repeated measures of the effects of ZEBO use on spelling for cohort 1 (2004)

Table 1 *Multilevel repeated measures analyses of conceptual use and instrumental use of ZEBO for spelling for cohort 1 (2004)*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	104.19 (1.43)	104.20 (1.25)
Time	4.43 (0.24)	4.43 (0.21)
Av. vs. low IQ	4.12 (0.44)	4.12 (0.44)
High vs. low IQ	9.78 (0.53)	9.77 (0.53)
Girls vs. Boys	2.24 (0.29)	2.24 (0.29)
Conceptual use in 2004	-0.17 (0.22)	
Time*conceptual use	0.01 (0.04)	
Instrumental use in 2004		-0.11 (0.12)
Time*instrumental use		0.00 (0.00)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	10.53 (2.13)	10.47 (2.12)
σ^2_{v1}	-1.42 (0.33)	-1.42 (0.33)
Level 2 (pupil)		
σ^2_{u0}	13.59 (1.31)	13.59 (1.31)
σ^2_{u1}	0.60 (0.22)	0.60 (0.20)
Level 3 (occasion)		
σ^2_e	21.22 (0.41)	21.22 (0.41)

Note: Standard errors between brackets, *means interaction effect.

Appendix 5.6 Multilevel analyses with repeated measures of the effects of ZEBO use on mathematics for cohort 1 (2004)

Table 1 *Multilevel repeated measures analyses of conceptual use and instrumental use of ZEBO for mathematics for cohort 1 (2004)*

	Model Conceptual use	Model Instrumental use
<i>Fixed</i>		
Intercept	46.16 (2.78)	45.05 (2.67)
Time	6.90 (0.37)	7.08 (0.32)
Av. vs. low IQ	7.17 (0.68)	7.17 (0.68)
High vs. low IQ	16.63 (0.82)	16.62 (0.82)
Girls vs. Boys	-2.19 (0.43)	-2.19 (0.43)
SES (1.25 vs. 1.0)	-2.17 (0.65)	-2.18 (0.65)
SES (1.9 vs. 1.0)	-6.43 (0.88)	-6.46 (0.88)
Nr of pupils in the classroom	0.22 (0.09)	0.21 (0.09)
Conceptual use in 2004	-0.16 (0.32)	
Time*conceptual use	-0.01 (0.06)	
Instrumental use in 2004		0.05 (0.18)
Time*instrumental use		-0.02 (0.03)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	21.23 (4.51)	21.47 (4.55)
σ^2_{v1}	-2.83 (0.71)	-2.84 (0.71)
Level 2 (pupil)		
σ^2_{u0}	51.36 (3.46)	51.35 (3.46)
σ^2_{u1}	-1.33 (0.44)	-1.32 (0.44)
Level 3 (occasion)		
σ^2_e	37.34 (0.73)	37.34 (0.73)

Note: Standard errors between brackets, * means interaction effect.

Appendix 5.7 Multilevel analyses with repeated measures of the effects of ZEBO use on spelling for cohort 2 (2004)

Table 1 *Multilevel repeated measures analyses of conceptual use and instrumental use of ZEBO for spelling for cohort 2 (2004)*

	Model Conceptual use	Model Instrumental use
<i>Fixed</i>		
Intercept	116.04 (0.71)	123.86 (0.38)
Time	3.61 (0.07)	3.62 (0.07)
Av. vs. low IQ	3.86 (0.45)	3.85 (0.45)
High vs. low IQ	9.32 (0.51)	9.31 (0.51)
Girls vs. Boys	2.20 (0.30)	2.18 (0.30)
Conceptual use in 2004	0.20 (0.42)	
Time*conceptual use	-0.04 (0.08)	
Instrumental use in 2004		-0.23 (0.42)
Time*instrumental use		-0.03 (0.08)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	7.47 (1.17)	7.33 (1.64)
σ^2_{v1}	-1.17 (0.27)	-1.18 (0.27)
Level 2 (pupil)		
σ^2_{u0}	19.50 (1.04)	19.52 (1.04)
σ^2_{u1}	0.00 (0.00)	0.00 (0.00)
Level 3 (occasion)		
σ^2_e	22.65 (0.42)	22.65 (0.42)

Note: Standard errors between brackets, * means interaction effect.

Appendix 5.8 Multilevel analyses of the effects of ZEBO use on mathematics for cohort 2 (2004)

Table 1 *Multilevel repeated measures analyses of conceptual use and instrumental use of ZEBO for mathematics for cohort 2 (2004)*

	Model Conceptual use	Model Instrumental use
<i>Fixed</i>		
Intercept	70.59 (1.57)	70.60 (1.57)
Time	6.03 (0.10)	6.02 (0.10)
Av. vs. low IQ	8.14 (0.62)	8.12 (0.62)
High vs. low IQ	16.39 (0.72)	16.38 (0.72)
Girls vs. Boys	-3.07 (0.41)	-3.09 (0.42)
SES (1.25 vs. 1.0)	-2.49 (0.57)	-2.48 (0.57)
SES (1.9 vs. 1.0)	-4.66 (0.88)	-4.73 (0.87)
Nr of pupils in the classroom	0.18 (0.05)	0.19 (0.05)
Conceptual use in 2004	-0.33 (0.54)	
Time*conceptual use	0.11 (0.11)	
Instrumental use in 2004		-0.71 (0.53)
Time*instrumental use		0.11 (0.11)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	12.03 (2.74)	11.62 (2.67)
σ^2_{v1}	-2.15 (0.50)	-2.10 (0.59)
Level 2 (pupil)		
σ^2_{u0}	36.58 (1.92)	36.59 (1.92)
σ^2_{u1}	0.00 (0.00)	0.00 (0.00)
Level 3 (occasion)		
σ^2_e	38.39 (0.71)	38.39 (0.71)

Note: Standard errors between brackets, * means interaction effect.

Appendix 5.9 Multilevel analyses of the effects of ZEBO use on spelling for cohort 1 (2006)

Table 1 *Multilevel repeated measures analyses of the effect of conceptual use and instrumental use on growth in spelling achievement for cohort 1 (2006)*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	100.62 (3.20)	99.35 (3.48)
Time	4.03 (0.34)	4.14 (0.37)
Av. vs. low IQ	4.61 (0.74)	4.61 (0.74)
High vs. low IQ	10.37 (0.86)	10.38 (0.86)
Girls vs. Boys	1.94 (0.45)	1.94 (0.45)
Number of pupils	0.20 (0.09)	0.21 (0.12)
SES (1.25 vs. 1.0)	0.12 (0.68)	0.12 (0.67)
SES (1.9 vs. 1.0)	4.53 (1.78)	4.55 (1.78)
Language at home (Dutch vs. Turkish)	-3.99 (1.89)	-3.99 (1.89)
Language at home (Dutch vs. Dialect)	0.81 (1.75)	0.83 (1.75)
Language at home (Dutch versus other)	-7.05 (2.50)	-7.04 (2.50)
Conceptual use in 2006	0.12 (0.29)	
Time*conceptual use	0.00 (0.05)	
Instrumental use in 2006		0.16 (0.19)
Time*instrumental use		-0.01 (0.03)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	5.88 (1.99)	5.75 (1.96)
σ^2_{v1}	-0.65 (0.27)	-0.64 (0.27)
Level 2 (pupil)		
σ^2_{u0}	18.11 (1.77)	18.10 (1.77)
σ^2_{u1}	0.21 (0.18)	0.21 (0.18)
Level 3 (occasion)		
σ^2_e	19.95 (0.52)	19.95 (0.52)

Note: Standard errors between brackets, * means interaction effect.

Appendix 5.10 Multilevel analyses of the effects of ZEBO use on mathematics for cohort 1 (2006)

Table 1 *Multilevel repeated measures analyses for mathematics cohort 1 for conceptual and instrumental use*

	Model conceptual use	Model instrumental use
<i>Fixed</i>		
Intercept	56.67 (4.35)	55.62 (4.68)
Time	5.92 (0.41)	5.77 (0.45)
Av. vs. low IQ	7.05 (1.06)	7.06 (1.06)
High vs. low IQ	14.72 (1.23)	14.72 (1.23)
Girls vs. Boys	-2.67 (0.65)	-2.69 (0.64)
Nr of pupils in the classroom	0.54 (0.12)	0.54 (0.13)
SES (1.25 vs. 1.0)	-2.42 (0.97)	-2.40 (0.97)
SES (1.9 vs. 1.0)	-4.12 (1.25)	-4.17 (1.25)
Combination class yes vs. no	-3.27 (1.29)	-3.20 (1.35)
Conceptual use in 2006	-0.48 (0.35)	
Time*conceptual use	-0.08 (0.06)	
Instrumental use in 2006		-0.22 (0.25)
Time*instrumental use		0.06 (0.04)
<i>Random</i>		
Level 3 (class)		
σ^2_{v0}	6.65 (2.81)	7.16 (2.95)
σ^2_{v1}	-0.37 (0.35)	-0.40 (0.36)
Level 2 (pupil)		
σ^2_{u0}	48.25 (4.21)	48.22 (4.21)
σ^2_{u1}	-1.14 (0.40)	-1.14 (0.40)
Level 3 (occasion)		
σ^2_e	35.63 (0.93)	35.63 (0.93)

Note: Standard errors between brackets, * means interaction effect.

Appendix 5.11 Perceived effects of the use of ZEB0: descriptive statistics

Table 1 *Effects of the use of ZEB0 reported by principals and teachers*

The use of ZEB0 leads to improvement in:	Respondents	N	Year	To a great degree		To a moderate degree		To a small degree		To a minimal degree/not		Missing/ I do not know	
				%	(n)	%	(n)	%	(n)	%	(n)	%	(n)
consultation on school functioning and quality	Principals	41	2003	5	(2)	34	(14)	24	(10)	34	(14)	2	(1)
		48	2004	2	(1)	38	(18)	33	(16)	23	(11)	4	(2)
		25	2006	12	(3)	40	(10)	36	(9)	12	(3)	0	(0)
	Teachers	220	2003	1	(3)	16	(35)	16	(35)	51	(113)	16	(34)
		236	2004	2	(4)	20	(47)	27	(63)	45	(106)	7	(16)
		141	2006	2	(3)	25	(35)	33	(47)	32	(45)	8	(11)
educational leadership	Principals	41	2003	0	(0)	12	(5)	46	(19)	32	(13)	10	(4)
		48	2004	0	(0)	15	(7)	29	(14)	42	(20)	15	(7)
		25	2006	0	(0)	20	(5)	32	(8)	32	(8)	16	(4)
	Teachers	220	2003	1	(2)	6	(13)	6	(14)	53	(117)	34	(74)
		236	2004	1	(2)	7	(17)	14	(32)	45	(107)	33	(78)
		141	2006	0	(0)	10	(14)	14	(19)	45	(63)	32	(45)
professional development	Principals	41	2003	2	(1)	24	(10)	37	(15)	32	(13)	5	(2)
		48	2004	0	(0)	25	(12)	31	(15)	40	(19)	4	(2)
		25	2006	4	(1)	24	(6)	36	(9)	36	(9)	0	(0)
	Teachers	220	2003	1	(1)	6	(13)	11	(24)	60	(130)	23	(51)
		236	2004	1	(2)	9	(22)	14	(33)	58	(137)	18	(42)
		141	2006	0	(0)	17	(24)	18	(25)	49	(69)	16	(23)
achievement orientation	Principals	41	2003	2	(1)	20	(8)	22	(9)	51	(20)	5	(2)
		48	2004	0	(0)	23	(11)	27	(13)	44	(21)	6	(3)
		25	2006	4	(1)	20	(5)	36	(9)	32	(8)	0	(0)
	Teachers	220	2003	1	(3)	11	(25)	11	(24)	55	(120)	22	(48)
		236	2004	1	(2)	9	(20)	14	(34)	59	(140)	17	(40)
		141	2006	0	(0)	17	(24)	21	(30)	45	(63)	17	(24)
team cohesion	Principals	41	2003	0	(0)	15	(6)	22	(9)	56	(23)	7	(3)
		48	2004	2	(1)	10	(5)	23	(11)	44	(21)	21	(10)
		25	2006	0	(0)	4	(1)	48	(12)	40	(10)	8	(2)
	Teachers	220	2003	1	(3)	6	(13)	11	(23)	64	(140)	19	(41)
		236	2004	1	(1)	7	(17)	13	(30)	62	(147)	17	(41)
		141	2006	1	(2)	11	(15)	12	(17)	58	(82)	18	(25)
pupil care	Principals	41	2003	5	(2)	15	(6)	20	(8)	59	(24)	2	(1)
		48	2004	6	(3)	19	(9)	21	(10)	46	(22)	8	(4)
		25	2006	4	(1)	8	(2)	44	(11)	36	(9)	8	(2)
	Teachers	220	2003	2	(5)	12	(27)	9	(19)	53	(117)	24	(52)
		236	2004	3	(7)	11	(27)	14	(32)	54	(128)	18	(42)
		141	2006	4	(5)	16	(22)	10	(14)	55	(77)	16	(23)
didactic methods	Principals	41	2003	0	(0)	7	(3)	10	(4)	56	(23)	27	(11)
		48	2004	0	(0)	10	(5)	29	(14)	54	(26)	6	(3)
		25	2006	0	(0)	8	(2)	44	(11)	20	(6)	24	(6)
	Teachers	220	2003	0	(0)	3	(7)	9	(19)	67	(148)	21	(46)
		236	2004	0	(0)	5	(11)	12	(29)	87	(204)	12	(29)
		141	2006	0	(0)	11	(15)	24	(34)	52	(73)	13	(19)
pupil achievement evaluation	Principals	41	2003	0	(0)	12	(5)	15	(6)	59	(24)	15	(6)
		48	2004	0	(0)	23	(11)	13	(6)	54	(26)	10	(5)
		25	2006	0	(0)	8	(2)	36	(9)	48	(12)	8	(2)
	Teachers	220	2003	1	(2)	8	(18)	10	(21)	70	(153)	12	(26)
		236	2004	2	(4)	6	(15)	11	(27)	68	(161)	12	(29)
		141	2006	1	(1)	9	(13)	24	(34)	59	(83)	7	(10)
adaptive education	Principals	41	2003	0	(0)	15	(6)	10	(4)	71	(29)	5	(2)
		48	2004	0	(0)	13	(6)	15	(7)	58	(28)	15	(7)
		25	2006	0	(0)	12	(3)	32	(8)	52	(13)	4	(1)
	Teachers	220	2003	2	(4)	11	(25)	7	(16)	62	(137)	17	(38)
		236	2004	1	(2)	9	(20)	11	(26)	63	(149)	17	(39)
		141	2006	1	(2)	11	(16)	21	(30)	54	(76)	12	(17)
Negative effects	Principals	41	2003	2	(1)	5	(2)	7	(3)	76	(31)	10	(4)
		48	2004	4	(2)	0	(0)	8	(4)	71	(34)	17	(8)
		25	2006	0	(0)	4	(1)	12	(3)	72	(18)	12	(3)
	Teachers	220	2003	4	(8)	1	(3)	5	(10)	66	(145)	25	(54)
		236	2004	2	(4)	2	(5)	2	(5)	72	(170)	22	(52)
		141	2006	0	(0)	1	(2)	5	(7)	75	(105)	19	(27)

Appendix 6.1 Correlations among predictors for the principal data

Table 1 *Correlations among the predictors of the use of ZEBO and the variables excluded from the regression model based on the principal data (2003)*

Predictor	Fit of output with user need	Monitors the quality of education	Experiments to improve education
Excluded variables			
Relevance of output	0.54**	0.26	0.10

Note: * $p < .05$, ** $p < .01$.

Table 2 *Correlations among the predictors of the conceptual and instrumental use of ZEBO and the variables excluded from the regression model based on the principal data (2004)*

Predictor	Fit of output with user need	ZEBO leads to quality improvement	Not afraid of changes	Principal encouragement professional development
Excluded variables				
Encouragement (to use ZEBO) by principal	0.46**	0.31*	-0.06	0.36*

Note: * $p < .05$, ** $p < .01$.

Table 3 *Correlations among the predictors of the conceptual and instrumental use of ZEBO and the variables excluded from the regression model based on the principal data (2006)*

Predictor	Clarity of goal	ZEBO use leads to quality improvement	Team decision
Excluded variables			
Encouragement (to use ZEBO) by principal	0.69**	0.67**	0.69**

Note: * $p < .05$, ** $p < .01$.

Appendix 6.2 Multilevel analyses on the teacher data in 2004

Table 1 *Variables influencing the conceptual use of ZEBO (2004)*

	Model 0		Model 1	
	<i>(N=120)</i>		<i>(N=120)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.
Intercept	-0.01	0.100	-0.05	0.09
Ease of data input			0.21	0.07
Satisfaction with amount of training			0.17	0.08
Encouragement by principal			0.19	0.08
Not afraid of changes			0.25	0.07
Time and resources			0.16	0.07
Teachers can influence ZEBO measures			0.23	0.07
Variance components				
Between schools	0.27	0.10	0.18	0.07
Between teachers	0.52	0.07	0.40	0.06
Percentage explained				
Between schools			27.40	
Between teachers			30.47	
Deviance	363		313	
Improvement in model fit (p)			0.00	

Table 2 *Variables influencing the instrumental use of ZEBO (2004)*

	Model 0		Model 1	
	<i>(N=85)</i>		<i>(N=85)</i>	
Fixed effects	Est.	S.E.	Est.	S.E.
Intercept	0.05	0.12	0.02	0.10
Time requirements of use			0.22	0.08
Encouragement by principal			0.23	0.09
Teachers can influence ZEBO measures			0.30	0.09
Principal encouragement professional development			0.26	0.09
Variance components				
Between schools	0.29	0.13	0.20	0.09
Between teachers	0.56	0.10	0.40	0.07
Percentage explained				
Between schools			30.24	
Between teachers			29.73	
Deviance	265		225	
Improvement in model fit (p)			0.00	

Appendix 6.3 Multilevel analyses on the teacher data in 2006

Table 1 Variables influencing the conceptual use of ZEBO

	Model 0 (N=101)		Model 1 (N=101)	
	Est.	S.E.	Est.	S.E.
Fixed effects				
Intercept	-0.04	0.15	-0.09	0.12
Time and resources			0.27	0.08
team cohesion			-0.37	0.11
Variance components				
Between schools			0.24	0.10
Between teachers			0.44	0.07
Percentage explained				
Between schools			24.95	
Between teachers			42.40	
Deviance	249		232	
Improvement in model fit (p)				0.01

Table 2 Variables influencing the instrumental use of ZEBO

	Model 0 (N=85)		Model 1 (N=85)	
	Est.	S.E.	Est.	S.E.
Fixed effects				
Intercept	-0.11	0.17	-0.09	0.14
ZEBO use leads to quality improvement			0.34	0.09
Time and resources			0.19	0.08
Teachers exchange information			0.19	0.10
Variance components				
Between schools			0.38	0.15
Between teachers			0.34	0.07
Percentage explained				
Between schools			28.59	
Between teachers			30.21	
Deviance	190		167	
Improvement in model fit (p)				0.00

Appendix 6.4 Correlations among predictors for the teacher data

Table 1 *Correlations among the predictors of the use of ZEBO and the variables excluded from the model based on the teacher data (2003)*

Predictor	Encouragement (to use ZEBO) by principal	Clarity of goal	ZEBO use leads to quality improvement	Team decision	Teachers influence ZEBO measures	ZEBO score
Excluded variables						
Fit of output with user need	0.29**	0.63**	0.52**	0.39**	0.43**	-0.15
Ease of data entry	0.26**	0.47**	0.46**	0.18*	0.44**	-0.03

Note: * $p < .05$, ** $p < .01$.

Table 2 *Correlations among the predictors of the use of ZEB0 and the variables excluded from the model based on the teacher data (2004)*

Predictors	Ease of data entry	Time requirements of use	Satisfaction with amount of training	Encouragement by principal	Not afraid of changes	Time and resources	Teachers can influence ZEB0 measures	Principal encouragement professional development
Excluded variables								
Accuracy of output	0.20*	0.37**	0.13	0.18*	-0.12	-0.07	0.15	0.09
Fit of output with user needs	0.31*	0.42**	0.05	0.43**	0.01	0.10	0.32**	0.37**
Hours of training and support received	0.03	0.08	0.31**	0.11	0.17*	0.31**	0.09	0.02
Satisfaction with amount of support	0.28**	0.16	0.89**	0.28**	0.08	0.20*	0.07	0.16
Clarity of goal	0.16*	0.23**	0.13	0.36**	-0.01	0.11	0.30**	0.27**
ZEB0 use leads to quality improvement	0.24**	0.06	0.15	0.42**	0.25**	0.22*	0.43**	0.19**
Team decision	0.11	0.23**	-0.01	0.15*	-0.06	-0.07	0.38**	0.38**
Monitors the quality of education	0.09	0.13	0.10	0.38**	-0.03	0.05	0.30**	0.53**
ZEB0 score	0.08	0.13	-0.02	0.22**	0.03	-0.09	0.11	0.11

Note: * $p < .05$, ** $p < .01$.

Table 3 *Correlations among the predictors of the conceptual and instrumental use of ZEB0 and the variables excluded from the regression model based on the teacher data (2006)*

Predictors	ZEB0 use leads to quality improvement	Time and resources	Teachers exchange information	Team cohesion
Excluded variables				
Fit of output with user needs	0.43**	0.21*	-0.10	0.02
Clarity of goal	0.29**	0.13	0.05	0.18*

Note: * $p < .05$, ** $p < .01$.

Appendix 6.5 Factors influencing the use of ZEBO: results from the interviews (2003)

In appendices 6.5 to 6.7 the results of the interviews are displayed. Behind the different variables one may find two numbers between brackets. The first number represents the number of respondents mentioning the variable. The second number represents the number of schools from which these respondents are a staff member.

Figure 1 Comparisons of the interviews with "LoSE", "AvSE" and "HiSE" of ZEBO with regard to school organisational characteristics (2003)

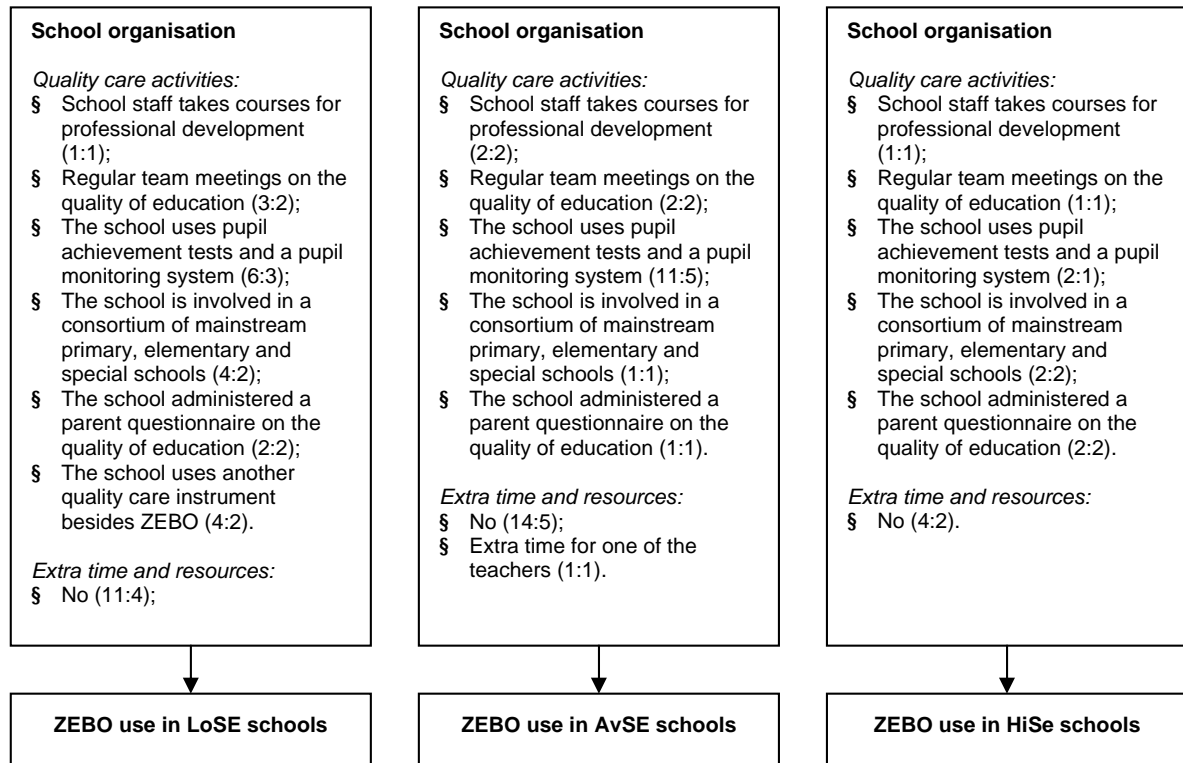


Figure 2 Comparisons of the interviews with “LoSE”, “AvSE” and “HiSE” of ZEBO with regard to characteristics of ZEBO (2003)

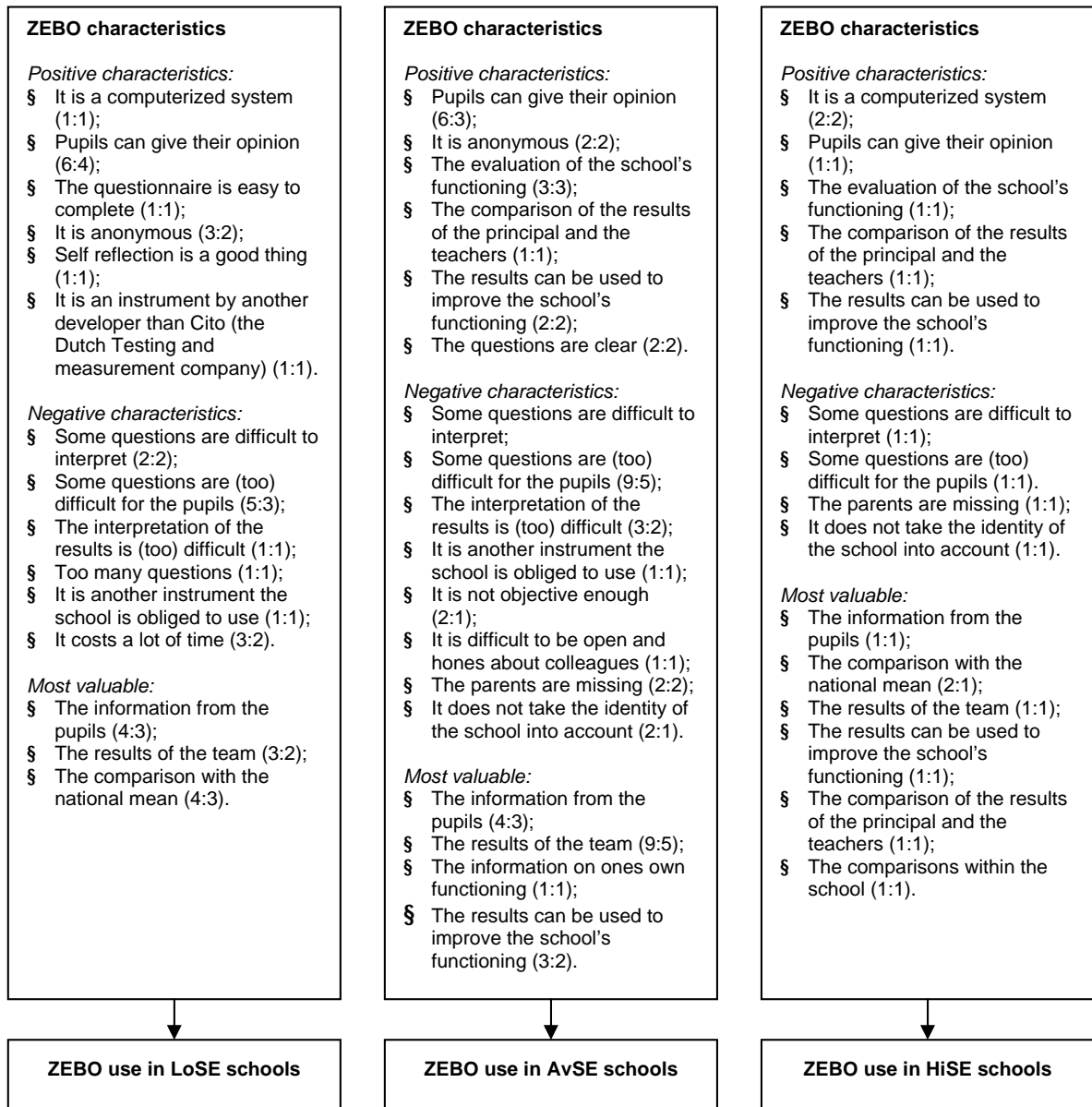
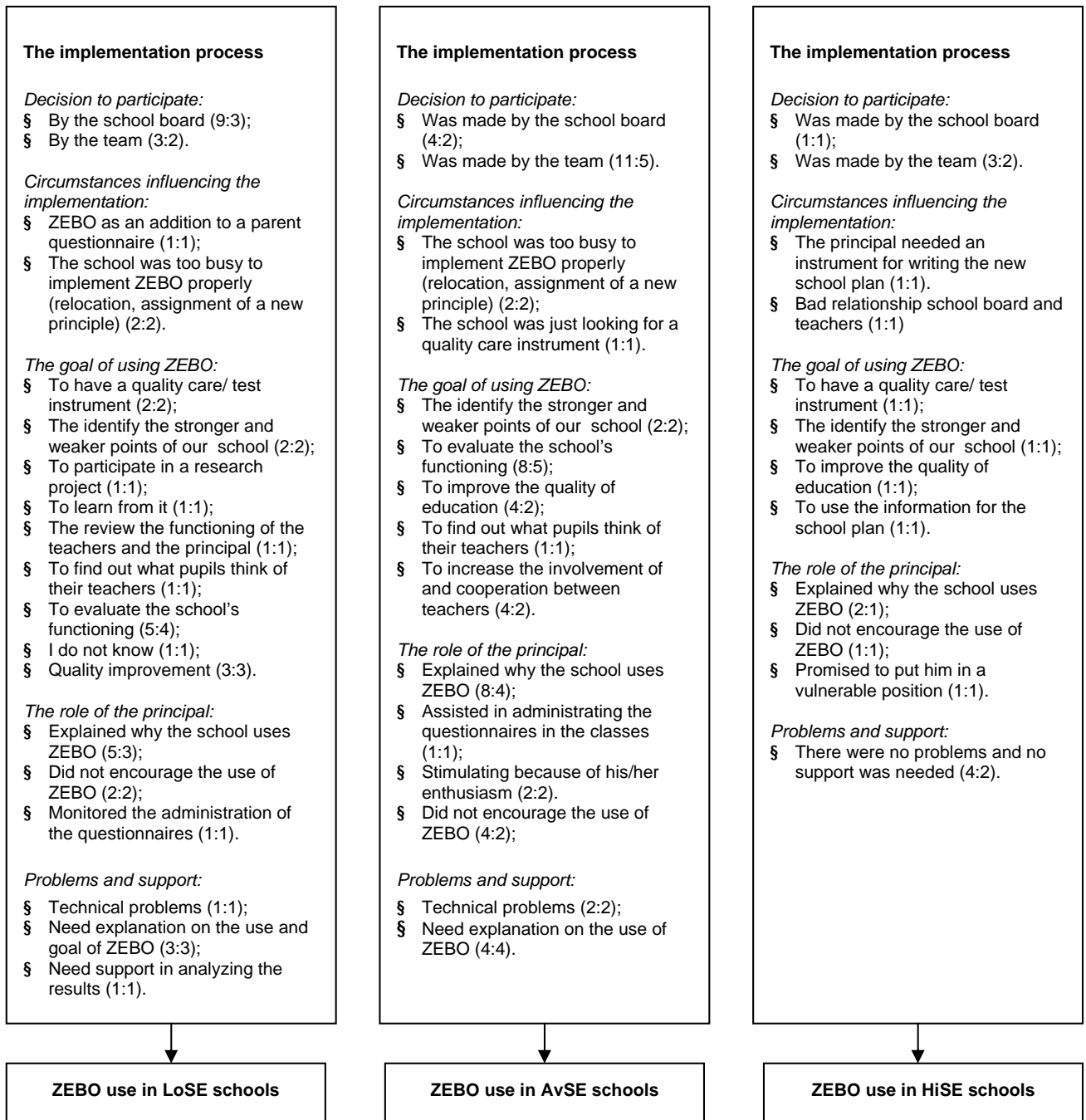


Figure 3 Comparisons of the interviews with “LoSE”, “AvSE” and “HiSE” schools with regard to the implementation process (2003)



Appendix 6.6 Factors influencing the use of ZEBO: results from the interviews (2005)

Figure 1 Comparisons of the interviews with “LoSE”, “AvSE” and “HiSE” schools with regard to school organisational characteristics (2005)

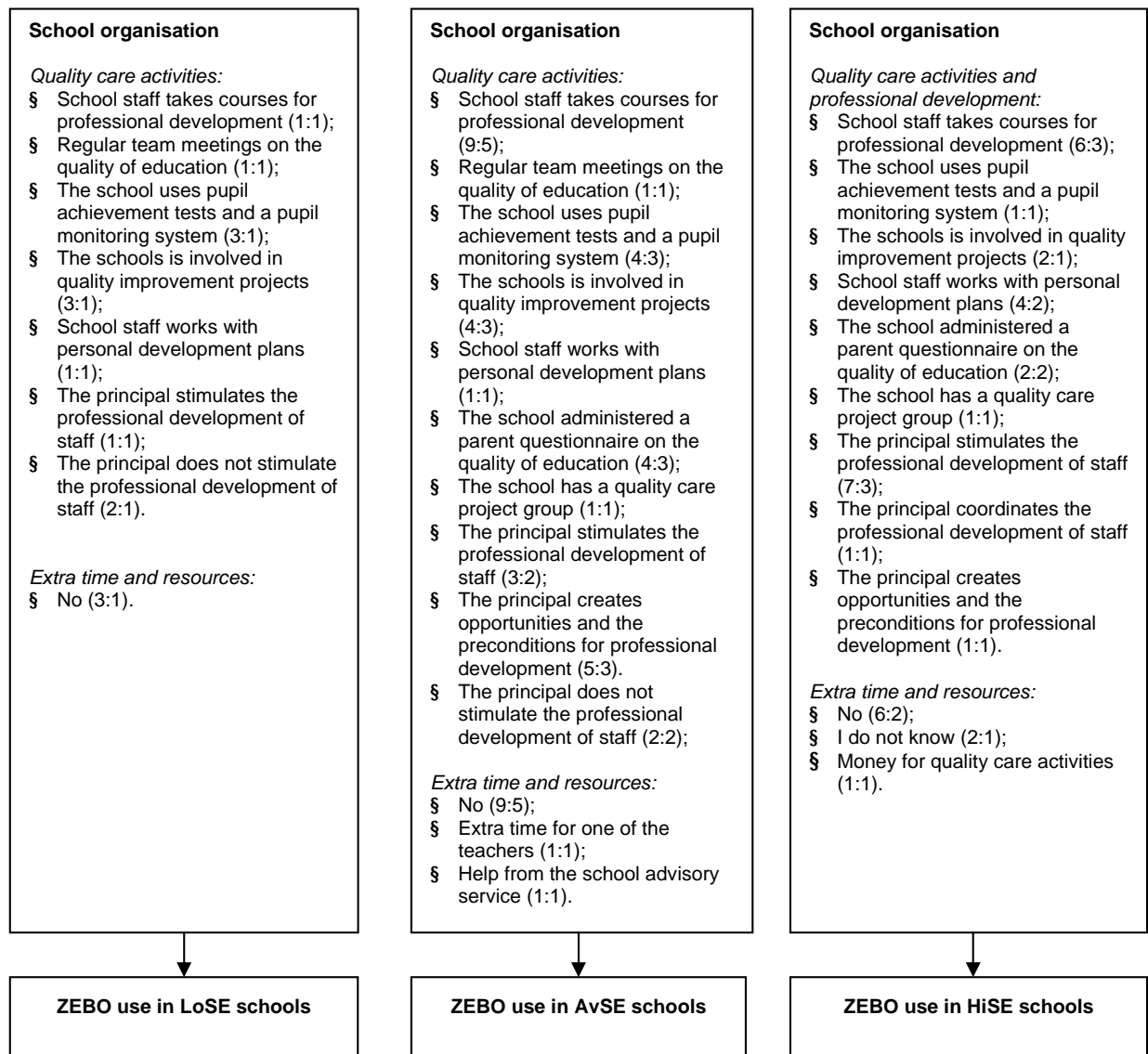


Figure 2 Comparisons of the interviews with “LoSE”, “AvSE” and “HiSE” schools with regard to the characteristics of ZEB0 (2005)

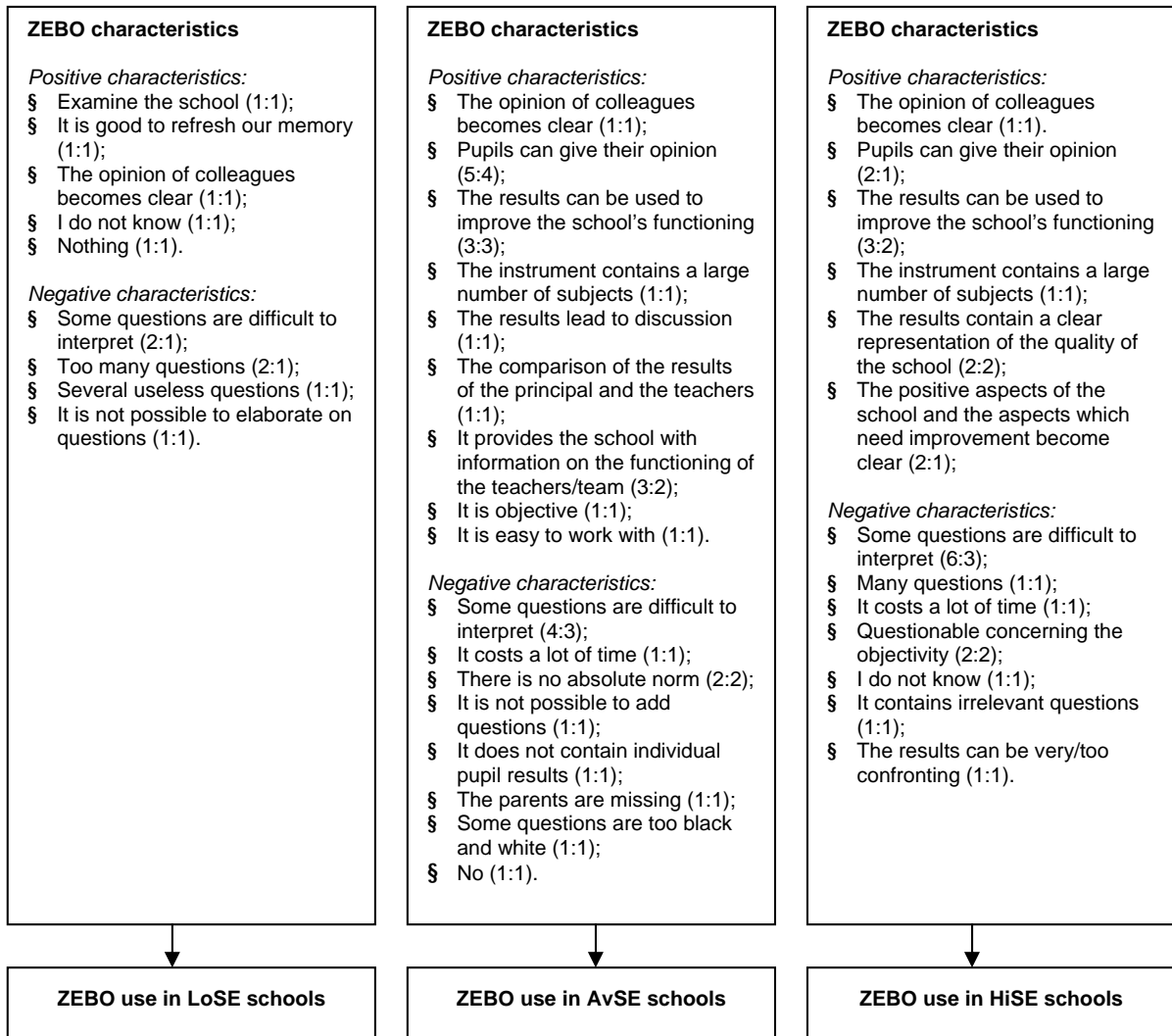
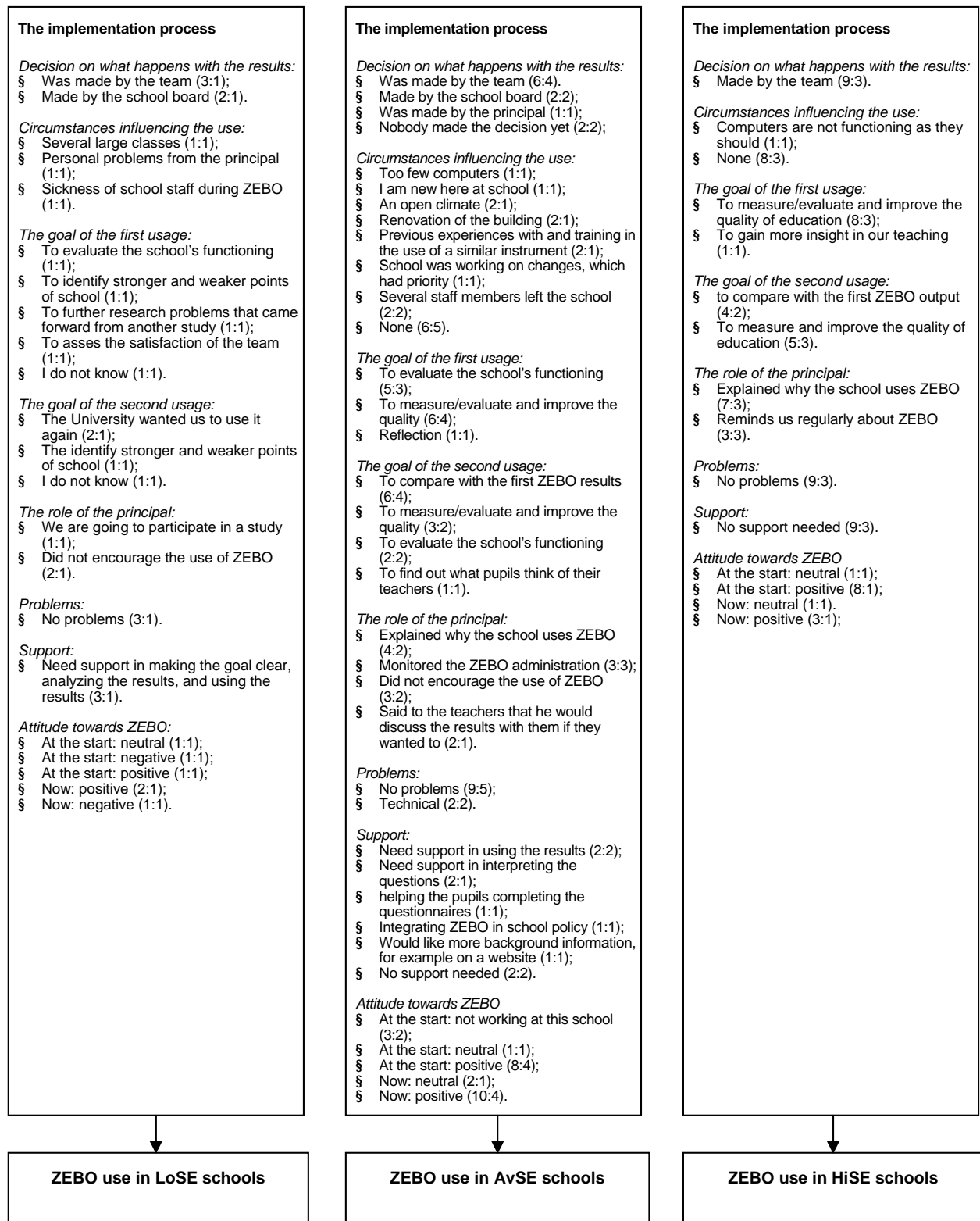


Figure 3 Comparisons of the interviews with “LoSE”, “AvSE” and “HiSE” schools with regard to the Implementation process (2005)



Appendix 6.7 Factors influencing the use of ZEBO: results from the interviews (2006)

Since only two principals were interviewed in 2006, no comparison could be made between the results of the LoSE, AvSE, and HiSE schools.

The two principals who were interviewed mentioned the following possible reasons for the lack of ZEBO use in several schools:

- § The principal has to encourage and sometimes even persuade school staff in using the ZEBO output;
- § Somebody has to take responsibility over the results;
- § An open climate in which issues are discussable is necessary;
- § The ZEBO instrument may not concur with the school staff's vision on education.