

This article was downloaded by: [Universiteit Gent]

On: 4 February 2010

Access details: Access Details: [subscription number 915540789]

Publisher Routledge

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



School Effectiveness and School Improvement

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t714592801>

Using school performance feedback: perceptions of primary school principals

Goedele Verhaeghe ^a; Jan Vanhoof ^b; Martin Valcke ^a; Peter Van Petegem ^b

^a Department of Educational Studies, Ghent University, Ghent, Belgium ^b Institute for Educational and Informational Sciences, University of Antwerp, Antwerp, Belgium

First published on: 06 November 2009

To cite this Article Verhaeghe, Goedele, Vanhoof, Jan, Valcke, Martin and Petegem, Peter Van(2009) 'Using school performance feedback: perceptions of primary school principals', School Effectiveness and School Improvement,, First published on: 06 November 2009 (iFirst)

To link to this Article: DOI: 10.1080/09243450903396005

URL: <http://dx.doi.org/10.1080/09243450903396005>

PLEASE SCROLL DOWN FOR ARTICLE

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

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

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

Using school performance feedback: perceptions of primary school principals

Goedele Verhaeghe^{a*}, Jan Vanhoof^b, Martin Valcke^a and Peter Van Petegem^b

^a*Department of Educational Studies, Ghent University, Ghent, Belgium;* ^b*Institute for Educational and Informational Sciences, University of Antwerp, Antwerp, Belgium*

(Received 28 October 2008; final version received 26 May 2009)

The present study focuses on the perception of primary school principals of school performance feedback (SPF) and of the actual use of this information. This study is part of a larger project which aims to develop a new school performance feedback system (SPFS). The study builds on an eclectic framework that integrates the literature on SPFSs. Through in-depth interviews with 16 school principals, 4 clusters of factors influencing school feedback use were identified: context, school and user, SPFS, and support. This study refines the description of feedback use in terms of phases and types of use and effects on school improvement. Although school performance feedback can be seen as an important instrument for school improvement, no systematic use of feedback by school principals was observed. This was partly explained by a lack of skills, time, and support.

Keywords: school performance feedback systems; school improvement; data use; qualitative research; user perceptions

Introduction

In recent years, the trend of decentralizing educational systems has prompted researchers to focus on school-based management and internal evaluation. Because schools are granted autonomy, governmental bodies expect them to be accountable for monitoring their internal quality policy (Nevo, 2002). In this context, the current performance level of a school serves as a starting point for developing future plans and educational targets. To assess their baseline performance level, schools can make use of feedback offered by school performance feedback systems (SPFSs). These external systems deliver confidential information about a school's performance and functioning (Visscher & Coe, 2002, 2003). Performance feedback helps to reveal the strengths and weaknesses of a school's functioning and is expected to contribute to the school improvement process by stimulating reflection and self-evaluation.

However, receiving feedback alone is not a sufficient condition to foster self-evaluation and systematic reflection at the school level. Several other conditions related to the school, the context, and the specific SPFS being used, determine if and how schools will make use of the available feedback. Empirical research on SPFSs is

*Corresponding author. Email: Goedele.Verhaeghe@UGent.be

limited (Schildkamp & Teddlie, 2008). Studies that have been carried out indicate that the actual use of school feedback and its impact are rather low (Coe, 2002; Saunders & Rudd, 1999; Tymms, 1995; Van Petegem & Vanhoof, 2004). We believe that a detailed study of the use and impact of existing school performance feedback initiatives is warranted (Goldstein & Spiegelhalter, 1996; Schildkamp, 2007; Schildkamp, Visscher, & Luyten, 2009; Visscher & Coe, 2002, 2003). In this study, we build on the findings of an ongoing project which focuses on the design, development, and implementation of a SPFS in Flanders (The Dutch-speaking community of Belgium). We investigate the school principals' perceptions of factors that promote or hinder their understanding and use of school performance feedback information. The results of this study are expected to support the development of SPFSs and to further refine theories on school feedback use.

Theoretical framework

Based on a literature review, we developed a conceptual framework that integrates factors affecting SPF use and effects (Fitz-Gibbon & Tymms, 2002; Schildkamp, 2007; Van Petegem & Vanhoof, 2007; Visscher, 2002; Visscher & Coe, 2003). This framework is presented in Figure 1.

School performance feedback use: phases, types, and effects

Adequate use of SPF is expected to lead to specific effects at the school and pupil level (Schildkamp, 2007; Visscher, 2002). Its purpose is to contribute to school improvement and lead to higher student performance (Visscher & Coe, 2003). Apart from the intended effects of SPF, unintended effects have also been reported in the literature, such as selective student admissions, teaching to the test, and removing difficult students (Visscher, 2002). Other studies refer to *undesirable side effects* of SPF, such as the demotivation of school staff who become overwhelmed by the amount of the data involved and the amount of time they have to invest (Fitz-Gibbon & Tymms, 2002; Schildkamp & Teddlie, 2008). In this context, SPF does not always result in significantly better student outcomes (Fitz-Gibbon & Tymms, 2002; Schildkamp et al., 2009; Visscher, 2002). Nevertheless, recent research indicates that SPF can have a positive impact on pupil achievement levels (Hammond &

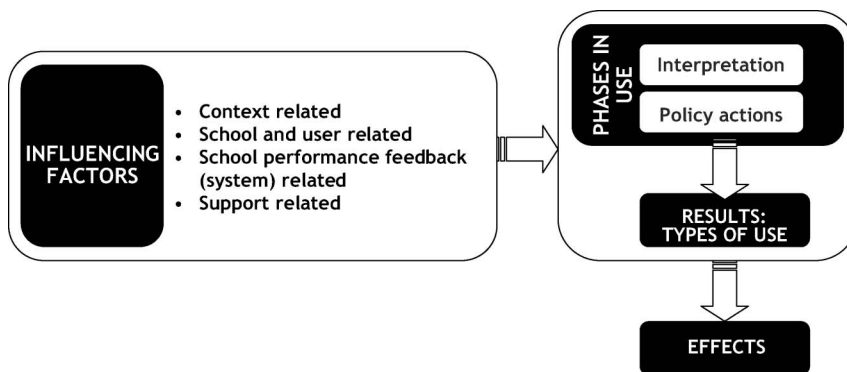


Figure 1. Conceptual framework of school performance feedback use.

Yeshanew, 2007) and on the associated school improvement processes (Schildkamp, 2007; Schildkamp & Teddlie, 2008; Schildkamp et al., 2009). In these studies, several effects on process indicators were observed, such as an improvement in consultation and communication about school functioning and school quality, improved didactical approaches, and a stronger achievement orientation of staff. However, considering the limited amount of research available, caution is warranted in drawing conclusions about the reported effects of SPF use (Coe, 2002; Schagen, 2004).

The way school feedback is used plays a key role in its potential impact. In terms of a policy-making cycle (e.g., Hoy & Miskel, 2001), feedback should be used in the following sequence. First, feedback results must reach the proper person(s). Second, the data in the report must be read and interpreted correctly for it to be meaningful. In the subsequent diagnostic process, causes and explanations for the results are deliberated. The diagnostic process results in actions that are implemented and finally evaluated. However, research indicates that school principals do not always disseminate feedback information or simply distribute feedback reports without examining them (Van Petegem & Vanhoof, 2004). Other studies found that school feedback users often get stuck in the transition from the interpretation of SPF to active policy-making (Schildkamp, 2007; Vanhoof, 2007). This is highly problematic as the interpretation of the data is essential in deducing workable information (Earl & Fullan, 2003). These phases of data use are outlined in the practice of data-driven decision-making (Learning Point Associates, 2004). However, in the current literature on SPF use for school improvement, these phases are not distinguished in a systematic way.

Within this policy-making cycle, different types of feedback use can be distinguished: (1) direct/instrumental, (2) conceptual, and (3) symbolic/convincing (Rossi & Freeman, 2004). An instrumental use of feedback serves as a starting point for immediate policy-making decisions. A conceptual use of feedback does not result in concrete actions but influences the decision-making process, which indirectly affects action. Even if feedback does not influence one's conceptualizations, it can affect the policy-making process in a symbolic way. This means feedback results serve to convince others of existing opinions and to support viewpoints in discussions (Visscher, 2002). Furthermore, feedback can be used in a strategic way for accountability purposes, although this is not in line with a school improvement discourse (Visscher & Coe, 2003). These four types of feedback use can be considered as results of feedback use. For example, a conceptual use results in an altered way of thinking about pupil performances. This intermediate result can in the end lead to effects of feedback use, such as a stronger achievement orientation.

Factors influencing school performance feedback utilization

Differences in the interpretation and use of school feedback can be attributed to a variety of factors. In the framework of Visscher (2002) and Visscher and Coe (2003), the following set of influential factors is outlined: context, school and user, SPFS, and support. The authors embed the process of feedback use in the broader school environment, which we call context-related factors. They do not distinguish support-related factors as a separate set but place them within the implementation process and characteristics of the feedback system. These variables were selected based on a literature review in the fields of educational innovation, educational management,

business administration, and computer science. However, the relations between the different influencing factors and the feedback effects are not examined (Visscher, 2002). This framework is used as a basis for the present study.

Context-related factors that impact feedback use include the school's policy strategies at the regional and/or governmental level (Sun, Creemers, & De Jong, 2007; Visscher, 2002). For instance, policies can contain clear expectations that schools make use of feedback information. Educational governments can stimulate feedback use by pressure and/or support. Furthermore, feedback will be used differently depending on the context (e.g., school improvement, school accountability, or a combination of both strategies) (Vanhoof & Van Petegem, 2007; Visscher, 2002).

Secondly, *school- and user-related* characteristics seem to be key variables explaining differences in school feedback use. First, the motivation to use a SPFS leads to different utilizations. Motivation varies from internal quality development or external accountability to policy preparation (Liket, 1992; Van Aanholt & Buis, 1990). Secondly, previous experiences with feedback use, general experience with school-related data, and the statistical knowledge and skills needed to interpret feedback reports will also influence feedback use. Whereas most teachers have experience with school test data, pupil monitoring systems, and self-evaluations, in several studies school staff report that they are lacking the skills and confidence when using data for school policy purposes (Earl & Fullan, 2003; Kerr, Marsh, Ikemoio, Darilek, & Barney, 2006; Saunders, 2000; Williams & Coles, 2007). Thirdly, school performance levels also influence feedback use (Visscher, 2002; Visscher & Coe, 2003). Schools receiving positive feedback (large value added) will discuss the results differently compared to schools receiving a less positive picture (Schildkamp, 2007). In line with control theory, participants receiving negative feedback are more likely to make an effort to reduce the discrepancy between the negative feedback and the expected standards (Kluger & DeNisi, 1996). This will result in different policy implications. However, this theory does not hold in all cases; it is not unusual for school principals to withhold feedback information that does not fit the current policy plan (Van Petegem & Vanhoof, 2004).

A third set of factors influencing school performance feedback use refers to the *characteristics of the school feedback reports and the feedback system*. In this context, the perception of the user determines how feedback will be used (Van den Berg & Ros, 1999; Visscher, 2002). At the level of content, feedback should be perceived as relevant, non-threatening, and corresponding to the actual informational needs (Schildkamp & Teddlie, 2008; Van Petegem & Vanhoof, 2007; Visscher, 2002). Furthermore, the representation of both absolute and relative school performance results also impacts the way feedback is used (Visscher, 2002; Visscher & Coe, 2003). If relative measures are used to compare the school's results with a reference group, these school scores should be adjusted for the influence of pupil background characteristics and should be linked to the relevant cohort group (Goldstein & Spiegelhalter, 1996). Information should also be up-to-date, reliable, and valid (Schildkamp & Teddlie, 2008; Visscher, 2002; Visscher & Coe, 2003). In terms of ethical issues, Fitz-Gibbon and Tymms (2002) refer to the Hippocratic Oath and state that feedback should "at least do no harm" (p. 75). For example, in some cases feedback can be threatening to the recipients' self-esteem, particularly in a system of accountability (Visscher & Coe, 2003). Consistent with our definition of SPFSs, feedback systems for school improvement should guarantee confidentiality and

anonymity to the subjects and schools. Moreover, feedback should not harm subjects or schools on the basis of misleading information (Goldstein & Myers, 1996).

The fourth and final set of factors that affect feedback use concerns the *support* experienced by feedback users (Schildkamp & Teddlie, 2008). School staff that are involved in SPFS training are more likely to read the feedback reports and adopt a more positive attitude (Tymms, 1995). Numerous studies stress the importance of providing feedback support (e.g., Schildkamp & Teddlie, 2008; Schildkamp et al., 2009; Van Petegem & Vanhoof, 2007; Visscher & Coe, 2003). This can be administered by educational and government parties, school team members, or the feedback system itself.

Research questions

This study examines the perception of school feedback users. Based on the conceptual framework discussed above, the following research questions are asked:

- (1) What phases can be observed in practice when schools use school performance feedback?
- (2) What is/are the result(s) of using school performance feedback?
- (3) How can differences be explained in the interpretation and the further use of school performance feedback in different school contexts?

Research context

This study is part of a larger SPF project called “Each school its own mirror”. As there is currently no SPFS available in Flanders, this project is in the process of developing and evaluating a new SPFS with collaboration between researchers, various stakeholders, and a target group of primary school principals and teachers. The system that has so far been developed from the SPF project gives schools feedback on a confidential basis. These feedback reports are designed to enable teachers and principals to understand the *value-added scores* of their school as compared to a reference group. The reference group used is taken from another research project (the SiBO project: Schoolloopbanen in het BasisOnderwijs [School Trajectories in Primary Education]) that is currently tracking approximately 6,000 children from a representative sample of Flemish schools from the time they entered kindergarten until the end of primary education. In the SPF project, scores on tests and survey and observational data are being continuously collected to gather information on child characteristics, family background, class characteristics, classroom practices, teacher attitudes and subjective theory, and school characteristics. The tests focus on language learning (orthography, reading fluency, reading comprehension) and mathematics. Item Response Theory (IRT)-based techniques are used to construct the test scores, enabling us to estimate growth curves.

The SPF project is currently able to deliver trial versions of school feedback reports to the 198 primary school principals participating. In this study, we build on the results from the trial versions sent to the schools in the spring of 2007. These reports inform schools about the performance of children and classes in the first 2 years of primary education. Results were reported for mathematics, reading fluency, and orthography, supplemented with information about pupil characteristics

(child factors, home factors, and Dutch language skills at the start of Grade 1). The school-specific results were compared to the Flemish reference group. The central concepts in these reports include *learning gain*, *value added*, and *adjusted scores* and were explained in such a way that no prior statistical knowledge was required. The data were supported with graphical representations (i.e., boxplots, bar graphs, pie graphs, growth curves, and cross tables). The text of each report was standardized. The school principals were required to interpret the results for their school, based on the general information made available. They also received individual pupil feedback, which represents the observed scores and percentile rankings relative to the reference group. Pupil feedback was presented to the schools shortly after taking the class tests, but the aggregated scores at class and school level were sent approximately 10 months later.

Research design

Research approach

In this study, we use a qualitative design to explore the perceptions of primary school principals on SPF use. A qualitative approach is appropriate since we want to develop a view on “naturally occurring, ordinary events in natural settings, so that we have a strong handle on what ‘real life’ is like” (Miles & Huberman, 1994, p. 10). It is recommended when the knowledge base is limited and the nature of the variables, processes, and interrelations is less clear (Maso & Smaling, 1998), which holds for the literature about SPF use.

Research instrument and procedure

Data were gathered on the basis of semistructured in-depth interviews. This type of interview creates an informal relationship between researcher and respondent and gives the researcher a better understanding of the perceptions, opinions, and views of respondents (Mason, 2002). The interview questions were largely open ended and were derived from the conceptual framework discussed above. Respondents were invited to describe their school situation, to propose suggestions, and to express their concerns. To clarify remarks or to ask for elaboration, spontaneous follow-up probes were allowed (Lindlof & Taylor, 2002). Examples of questions include:

- Questions about feedback characteristics: These questions focused on the perceptions of the relevance, interpretability, user-friendliness, validity, and reliability of the feedback information (e.g., Do you think the information is relevant to draw a picture of the school’s influence on pupils’ performances? Which information is the most relevant? Why? Do you trust the quality of the feedback results?).
- Questions about school and user characteristics mainly focused on interpretation skills, expectations of feedback use, and the perception of the school’s performance (e.g., Do you feel comfortable interpreting these feedback results? If yes, where did you acquire the knowledge and skills for this? Which problems did you encounter?) Furthermore, questions regarding school culture characteristics were asked (e.g., Is there a culture of systematic reflection? To what degree do teachers welcome school performance feedback reports?)

Besides this feedback project, are there other data-gathering systems used to assess the school's functioning?).

- Questions about support initiatives included support use and support needs (e.g., Have you engaged the team members when interpreting the feedback reports? Do you feel enough support from the school staff when interpreting the results? Is there a need for more external support? For which activities?).
- Questions on feedback use were formulated to discern different types and phases of feedback use (e.g., Did you formulate any goals you want to achieve by using feedback? Which initiatives are you undertaking to communicate the feedback results to staff members? Did the feedback report play a role in policy decision-making? Has it influenced your way of thinking about the school? Did you use the report for strategic purposes, such as promoting your school, informing the school inspectorate about your school's results? Did you use the report to legitimize your own convictions?).
- Questions about feedback effects were not stressed because it was unlikely that effects of feedback use on the school could already have been observed in the 3-month period between the feedback delivery and the interview. However, questions about participants' expectations of effects were posed (e.g., What effects should take place for your effort to have been worthwhile?).
- The perception of context-related factors is limited in this study to the influence of the inspection visits to schools.

School principals were visited in their school office by one of the two interviewers, 3 months after receiving their school performance feedback report. Interviews lasted approximately 90 min.

Theoretical sampling

From the 198 SiBO principals, a sample of 16 primary school principals was selected by means of theoretical sampling, maximizing a variety of feedback use (Mason, 2002; Silverman, 2005). In this sampling method, the choice of cases is made on conceptual grounds, not on representative grounds (Miles & Huberman, 1994). To gather this sample, 2 months after having received feedback reports, the 198 principals were asked to fill out an online survey. We obtained a response rate of 61%. The principals were selected for the present study on the basis of the following variables: the degree to which they used the school feedback, the number of children without special needs in their school, experience in working with self-evaluation, and school performance as represented in the feedback report. For each variable, the schools were divided into three groups (low, average, and high), with the exception of the school performance level (positive or negative value added). In this survey, the principals were asked with whom they had discussed the feedback report, and they chose from six answers. This was considered as an indicator of feedback use. Respondents that depicted more than three options were defined as *high users*. Principals that marked less than two options were defined as *low feedback users* ($M = 1.77$, $SD = 1.26$). The second variable concerns the school's performance level (Visscher, 2002; Visscher & Coe, 2003). A distinction was made between schools with a positive or negative value-added mathematics score at the end of Grade 2. In the online survey, principals were asked to report their degree of experience in conducting self-evaluations in the school. Respondents with scores higher than 3 on

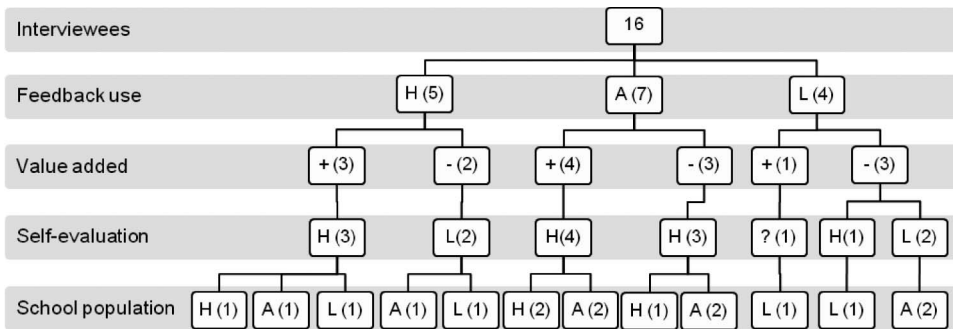


Figure 2. Overview of selected respondents.

Note: H = high, A = average, L = low, ? = information unknown, + = positive value-added score, - = negative value-added score; number of respondents between parentheses. From left to right, respectively, respondents from School 2, 11, 7, 1, 16, 3 & 10, 4 & 13, 15, 14 & 5, 6, 9, 8 & 12.

a 5-point Likert scale were classified as *highly experienced* and those with scores less than 3 as having a *low degree of experience* ($M = 3.50$, $SD = 1.08$). This selection criterion was used as it indicates prior experience in data use for school improvement. The fourth selection variable was the number of pupils without special educational needs (SEN) at their school. As feedback reports in this case were adjusted for pupil background characteristics, a differential approval of the feedback relevance was expected. Schools with percentages between 30 and 70 are considered as having an *average number of pupils without SEN* ($M = 50.36$, $SD = 27.73$). Figure 2 gives an overview of the selected schools.

Framework analysis

Next to influencing the design of the SPFS, the results of this study were also used as a means to evaluate the theoretical framework presented above. Therefore, the interview data were placed in the theoretical frame to examine whether the theoretical findings were confirmed or needed to be altered and/or elaborated. This can inspire future studies that build on new preliminary concepts and hypotheses (Ritchie & Spencer, 1994). These findings can also contribute to the ecological validity of research findings on feedback effects, as here they are applied in the context of school improvement (Visscher & Coe, 2003).

Each interview was transcribed verbatim and was independently coded by two researchers with ATLAS.ti. Codes were assigned by following the *middle order approach*, which allows for the initial application of broad categories that can later be refined (Dey, 1993). Text fragments were mainly assigned to codes in a deductive way. First, text fragments were placed under broad categories (e.g., effects of use, phases of use, the four groups of influencing factors, types of use, and other relevant information) and were then assigned to a predefined coding structure. If no predefined code was appropriate, the text fragments considered to be of importance were placed under the suitable broader category. New codes were created for these fragments inductively, emerging from the data, as in the grounded theory approach (Strauss & Corbin, 2007).

For inter-rater agreement, the first two interviews were coded collaboratively and the coding structure was set up. Two interviews were then coded by both researchers separately to calculate inter-rater reliability, following the formula of Miles and Huberman (1994): ratio between the number of agreements and the total number of attributed codes. An inter-rater correlation value of .90 was calculated, indicating good inter-rater reliability.

After this coding phase, the analysis shifted from a focus on individual interviews in a vertical analysis to a focus on the coding categories as they occurred in all the different interviews in a horizontal analysis (variable-oriented approach; Miles & Huberman, 1994). This allows the researcher to transcend the individual narratives of the school principals and to create a spectrum of perceptions and interpretations.

Findings and discussion

What phases can be observed when schools use school performance feedback?

The interview results confirm that school performance feedback use in primary schools is limited. Most schools were situated in the first phase of the policy cycle described above. Only a few schools reached the *planning phase* and *action phase* in the policy cycle.

Concerning the dissemination of information, the first stumbling block occurred at the moment feedback reports arrived at the school. Though all interviewees confirmed receipt of the report, one of them could not remember it. This stumbling block became more apparent when we examined the various ways in which the reports were handled. In some schools, the report was not read: "Mostly the reports arrive at the school. I give it a glimpse and then it is classified. Then, nothing is done with it" (School 8). Other school principals reported they only took a quick look at it. In contrast, others distributed the report to the teachers responsible for the class that was discussed in the report. Others handed the report over to the special needs teachers or special care coordinators. Sometimes teachers were intentionally not asked to be involved in reading the reports.

My opinion is that if you are not really acquainted with the interpretation of these data, you will not spontaneously unravel the whole report. It is not so easy. It is an extra task on top of the rest. If I do this and draw the conclusions and give it to them, it is already a lot. (School 5)

Occasionally, reading the feedback reports led to discussion between the principal and the special care coordinator. In other cases, teachers were also invited into a discussion, but even then it was not guaranteed that they would read the reports. Principals reported that informal and unplanned discussions took place:

We have a smoking room. That's where we discussed the report. Those who entered the room glanced through the report. It was not intentionally communicated to the rest of the team members. This happened rather informally. (School 10)

Other principals reported having a formal discussion during planned multi-disciplinary team meetings. In these instances, the school principal or special care coordinator presented a summary of the results and their interpretations. All school principals reported that they only discussed the feedback information within the school team, with the exception of also reporting the information to the education inspectorate.

While we made a theoretical distinction between a reading and an interpretation phase, it became clear that in practice these phases merged together. The principals or special care coordinators that discussed the results with team members proposed their own interpretations.

The new report was read and discussed by me and the care coordinator. Afterwards the report was discussed in a team meeting with all teachers; not just the teachers that are involved in the research. Conclusions and underlying statistical procedures were communicated. Growth curves were presented. (School 2)

Principals also stressed that the interpretation process was an intensive, time-consuming, and difficult activity. Some confirmed that they were not able to correctly interpret or understand the information. This is problematic as the interpretation phase is crucial for developing a solid and valid basis for the development of school policy (Earl & Fullan, 2003). While a minority reported not having experienced difficulties, all principals reported that successfully interpreting the report requires effort.

You really have to examine it carefully to figure it out. I went over it ... but to really master it, you have to read and examine it several times. (School 1)

I think that ... one of the reasons is that you first look at it. It is similar to the directions for use of a new apparatus. First you set it up and afterwards you read how it works. If the set up is successful, you are not going to read the instructions for use. (School 14)

The laborious interpretation phase seems to have a strong impact on the diagnostic phase. Most principals dropped out after one attempt at understanding the feedback results. Only a few principals set up initiatives to identify strengths and weaknesses in their school and examined the feedback information when looking for explanations. However, this was rarely set up in a systematic way.

Principals frequently stated that the diagnostic and action phase were barely reached. They also linked this to the lack of cues in the feedback reports that might direct future action. This may be a reason why school feedback is not systematically taken into consideration when developing internal policy.

We discuss it with the teachers involved. And, until now, the interpretation is limited to the reading of the report and the file, but no immediate actions follow from this. (School 4)

But this feedback is not that useful for classes and individual children. I think this is the biggest concern. In fact it has to be as concrete as possible. That is the request of teachers; something ready-made. In fact this is also partly how I am. If I take a method book, I expect not to have to search for accompanying exercises. (School 3)

What is/are the result(s) of using school performance feedback?

The findings discussed above indicate factors that can affect the outcomes of SPF use. We found that in some schools feedback is used as a mirror image of the school's performance. In those cases, a better understanding of the school's impact on pupil performance was developed. However, this did not automatically lead to (policy) actions. This can be labeled as *conceptual* feedback use; it led to reflection in schools, even when the results confirm prior findings and impressions.

Indeed, so far we have ... already noticed a few things concerning the school's position that we were not aware of before. What we also notice is that there is a large pupil

mobility, which influences our results significantly. These are important findings for us. (School 12)

Most important was to see where the school's position is. How well are we performing and whether the school realizes a value-added score. This is, for me personally, a refinement in thinking about what you are doing as a school, about your task, about your aims ... (School 7)

Illustrations of *instrumental* feedback usage were rare. Some principals stated that the feedback information did not offer enough starting points (e.g., remedial information) to direct actions. However, some principals reported that action had been taken, such as a reorganization of rosters, an increase in the number of teaching hours, the introduction of a new reading method, and more intensive mentoring of new teachers. Even when information confirmed prior findings, it led to instrumental feedback use: "What is reported confirms what we already assumed. It is more like an affirmation of our feelings. And we have done a few things, such as introducing a new spelling method" (School 10).

Feedback information was particularly used in a *symbolic* way. Respondents indicated that school feedback was a useful instrument in highlighting existing opinions and underlining various problems in the school's functioning. According to the respondents, the feedback was used as input for shared decision-making. However, this did not lead to concrete action.

I had my own vision of the school and I wanted to impose it on the team; this was a good instrument to make out a case for it and to say it is necessary that we deal with this. (School 4)

Examples that we found of *strategic* utilization referred to the use of school feedback in the development of the self-evaluation report to be submitted to the education inspectorate. Principals reported that they were grateful to participate in the study because they could make use of school-, class-, and pupil-related information for this purpose. This factor deviates from the original theoretical model of SPF usage. Schools seem to have used the feedback information in the context of being accountable to the inspection authorities. This is in contrast to the perception of the authors and developers of the SPFS, who want feedback to be used for school improvement.

Not all of the information gathered about feedback use could be placed within the predefined coding scheme that was based on the literature (Rossi & Freeman, 2004; Visscher, 2002). Therefore, two extra codes were created: a motivating use and a pupil-directed use. In some cases, the feedback information helped to *motivate or stimulate* school team members. In some schools, the feedback was communicated to team members for this purpose, which sometimes implied a selective presentation of the results.

If you are an immigrant school, as we are, sometimes it is questioned if our performance level is high enough. And if you receive an output report from an external organization, it partly confirms we are doing a good job. (School 16)

For making internal plans ... we selected some results for reading and mathematics. We used these results for our own reports to say: "Look, on this measurement occasion, we just took out these results and notice that our children score like this. And the Flemish average is like this. Thus, we are below this average." (School 7)

The latter statement illustrates that lower performance results were also used to motivate the team members to overcome shortcomings. Conversely, some school

principals kept the feedback results private, especially if they were not as good as expected. This was explained by the intention not to discourage team members.

For example, concerning the learning gain scores. Absolutely. If I had to communicate it and mention that for example the learning gain in the first grade is smaller than on average and in the second grade larger than on average, this would be very hard to bear for the teacher involved if this is made public. I am sure of that. (School 6)

All of the aforementioned examples indicate feedback usage at the school level. During the interviews, principals stressed that aggregated results were useful for policy-makers but not for teachers who prefer a *pupil-directed* utilization. Classroom teachers need data at the pupil level to direct actions that correspond to the learning needs of individual pupils. Pupil feedback is seen as complementary to pupil monitoring systems and is also considered more accessible to interpretation and to direct action on short notice.

These interview results indicate that school feedback is not extensively used and has a limited impact. In fact, many school principals had not yet noticed school improvement effects by using the SPF, and if they had, they referred to the effects of using the feedback reports of the previous year.

[As a result of mentoring starting teachers and introducing a new method; cf. instrumental use] We see the AVI results [AVI is a Dutch grading system for reading fluency often used in primary education]. When before almost no pupil reached an AVI-1 level at the end of the year with that method and that young teacher, we now have several AVI-6 levels. Thus we have good results. That partly was a result of that. (School 1)

Some principals stated that, because of the longitudinal nature of the study that provides the feedback services, barriers against the feedback discussions in the group decreased and interest in the results increased. This illustrates the valuable effects of process variables that indirectly contribute to school improvement (Schildkamp, 2007; Schildkamp & Teddlie, 2008; Schildkamp et al., 2009).

How can differences be explained in the interpretation and the use of school performance feedback in different school contexts?

In the theoretical framework, different factors/conditions were discerned that explain differences in school feedback use. Our findings confirm the distinction of four clusters of related factors.

Context-related factors

To understand why school feedback is used to such a limited extent, we must take into account both the research context and the Flemish educational context. In terms of the research context, the SPF presented information at the school level with adjusted scores. These built on a comparison with a reference group, resulting in value-added scores. This is a very new approach that principals are not acquainted with.

In terms of the Flemish educational context, the central educational authorities do not formally encourage or oblige schools to adopt a SPFS approach. Indeed, some authorities are even reluctant to do so, stating that it introduces the risk that

schools will be compared and ranked on the basis of biased information or that adjusted scores will reveal another school performance level than expected. However, educational inspection authorities adopt another view. They encourage schools to document school performance on the basis of performance-related information.

[On being questioned about whether it was a conscious choice to participate in the research project] You always have the possibility to refuse . . . The main reason for me to participate was that our inspectorate often asks for output results. And yes, of course we have our own class tests, but there is no reference point, because teachers create their own tests. We also have tests from our methods. But nowhere is there a comparison with another school to see how we perform. (School 1)

School- and user-related factors

The interview analysis indicated four groups of related school and user characteristics.

Differences in expected functions and effects of school performance feedback. School principals differed in the degree to which they had expectations of using feedback as well as in the goals they oriented themselves towards with feedback use. Some did not even define goals or targets, whereas others reacted in a proactive way. When schools did formulate explicit and shared goals, the chances of observing more optimal and successful feedback use increased. This indicates that, if schools are convinced of the potential of school feedback, they undertake actions toward effective use (Bosker, Branderhorst, & Visscher, 2007). These actions have to be performed by the users themselves for innovations to become successful (Fullan, 2007).

A distinction can be made between *utilization expectations* and *effect-oriented expectations*. In the former situation, school principals expected to use the school feedback as a mirror, helping to develop a clearer view of the current school operation and school performance, and to detect strengths and weaknesses. Others expected to use feedback for policy development (e.g., for evaluating policy decisions or developing policy plans).

We thought, “look, this research will be conducted over 7 years; we are going to follow it up. Where are we as a school? We are putting a lot of effort into our care policy. What does this effort give us in return?” . . . In fact, we do have a very problematic population and it is our goal to see what the benefit is of all our effort. (School 1)

Another utilization-oriented perspective was discussed above (i.e., when principals used the information for accountability purposes). Almost all principals intended to use the feedback as input for their discussions with inspection authorities but stressed that they would not do this for parents.

In terms of effect expectations, principals expected that investing time and effort in school feedback would eventually improve education: “We expect to improve our quality of education. So far, for the first grade, it was worth the effort. That is the goal: an improvement of our education” (School 1). We found no evidence that the principals systematically reflected upon their expectations with regard to feedback use and feedback effects. In addition, principals indicated that their expectations of the feedback did not necessarily reflect the opinions of their staff members.

Teachers are not willing to participate because it is a lot of work for them. Moreover, the SPF project examines the same domains as the pupil monitoring system, thus it does not directly benefit them. . . . Teachers participate in this research project because the previous school principal decided they would. For them, it is “if it must be.” (School 2)

Differences in statistical knowledge and skills. Most school principals claimed not to have advanced statistical knowledge. Their statistical knowledge was acquired during their initial teacher training and additional training courses and was partly based on learning to work with pupil monitoring systems. However, they stressed that this was insufficient to work with school performance feedback. Conversely, some did not experience difficulties, either because everything was explained in the report or because they had sufficient prior knowledge.

Everything [in the feedback reports] is explained in terms of how to interpret it. Thus, if one pays enough attention to the instructions “to read it this way and these numbers, if this is mentioned it means this”, then I think no extra prior knowledge is needed. (School 4)

Differences in time available for feedback use. Some principals reported that if more time was available, they would have made more use of the feedback. Because principals and teachers have to divide their time over a large number of activities, less urgent tasks as those related to SPF use are not prioritized. This confirms previous findings that the self-evaluation of a school is not a priority for principals and teachers (Visscher, 1996; Williams & Coles, 2007).

There is often a lack of time. You cannot use this as an excuse, but it is often the reason. For example at team meetings, you want to put this and this on the agenda, but then there is not enough time to go more deeply into it, because there are so many issues coming from the outside. (School 11)

Differences in perceptions of positive/negative feedback results. When school feedback reflected low performance levels, the principals were willing to search for explanations. This confirms the control theory of Kluger and DeNisi (1996). However, this observation cannot be generalized: When the performance levels were far below average, sometimes feedback results were not distributed in order not to discourage team members. When performance results were perceived to be relatively good, further use of the feedback reports decreased: “We are scoring on average, so there are no severe differences. So why should we pay much attention to it?” (School 3).

The perception of the performance results was influenced by the way the results were represented, for example, by the way the value added is calculated. The feedback reports presented both adjusted scores that took into account the influence of pupil background characteristics and nonadjusted scores. Our results indicate that, especially in schools with a large number of children with special needs, the adjusted performance scores were valued positively.

The surplus value of this research for our school is that for all these years we’ve had the impression we were doing things right. Because we have a large number of foreign-speaking and special needs children, we want to know the effects of the way we organize our education and monitor our children. . . . Particularly in the last few years with the introduction of adjusted scores, some attention is given to the pupils’ progress, while taking into account certain factors. (School 7)

School performance feedback (system)-related factors

Feedback has to meet a number of requirements to facilitate correct interpretation and to promote feedback utilization.

Differences in perceived feedback relevance. All school principals requested that feedback should fit their needs. These needs differed between schools. Some principals expressed a primary interest in performance results on mathematics and language; others were more interested in socioemotional development or other subjects. Furthermore, schools' preferences differed in the calculation of value-added scores (observed or adjusted scores), in the way information was aggregated (pupil – class – school – other subgroups), in the amount of statistical background information in the reports, and the nature of the reference group(s). During the interviews, these differences were observed between and within schools. Differences were also related to the roles and occupations of feedback users. Teachers prefer pupil-level feedback, pupil-relevant error analysis, and remedial material, whilst policy-makers prefer aggregated information that reflects their school focus.

In my opinion, the school and class level is the most interesting, in view of my function. I am supposed to work mainly on school and class level and less on the pupil level. Thus for me this is more interesting than an individual report. But of course a teacher will see it differently. I am sure of that. This teacher will probably prefer feedback about the pupils in this class. (School 10)

When asked for ideas on how to better meet user needs, respondents suggested enlarging the amount of school subjects to be tested, focusing on different pupil cohorts, and tailoring information. The interviewees were not pleased about redundant information. They required feedback systems to focus on complementary information. In particular, some principals asked for information that would complement the available monitoring systems. All respondents required that the performance feedback be up-to-date. In particular, teachers expected feedback within the same school year as when tests had been taken, in order to support low-scoring pupils. When teachers shifted classes, feedback results of previous years were considered irrelevant.

Differences in perceived feedback interpretability. For this factor, no coherent picture could be deduced from the interview data. Most principals stated that interpreting the information was difficult. Some stressed that interpreting the information without support was a hopeless task. Some stated that the information could not be understood after only one reading. But not all principals considered this to be a problem or experienced difficulty in analyzing the reports. Some principals stated that it is important to stress that school feedback is a complex field and cannot be simplified without losing depth and meaning.

It is magnificent the way this report [is written] . . . It is not easy to explain something complex that clearly. They [the feedback developers] largely succeeded in it, but it is still a large amount of information. . . . Of course, sometimes I get lost, which is not surprising, considering the technicality of it. (School 7)

During the interviews, explanations for why some principals were not able to correctly interpret the feedback were given. Some complained of a lack of structure

in the feedback information. Others criticized the amount, stating that they skipped a lot of information, were selective, and focused only on the school results.

Maybe some parts are less interesting for me, but this is not a reason to leave out this information from the reports, because everything is concisely described. For example, the information about pupil mobility, if it does not interest you, just turn the page. (School 10)

In contrast, others appreciated the comprehensiveness of the feedback reports and preferred the additional information. A third element influencing the interpretability of feedback was the balance between technical concepts and the way school staff label and discuss education. Feedback was often experienced as being too abstract. Additionally, principals seemed to be less familiar with feedback that was aggregated at the class and school level. Both the language used and the graphical representations (growth curves, box plots) led to difficulties in interpretation. Some school principals stressed that the feedback is not appropriate for teachers as they do not possess the competence or experience to interpret the information, whilst others did not question the competence of their staff.

Differences in perceived validity and reliability. Respondents trusted the professionalism of the feedback developers. Nevertheless, they expressed some concerns. Some principals valued the feedback less because the adjusted scores do not take into account school-specific process and context variables. The feedback developers wanted to articulate these differences, but schools preferred an adjustment model taking into account more external influences that explain school outcomes and result in an average school profile.

I think researchers do not have enough information [about pupil and school characteristics]. They do not know we introduced a new reading method, which caused problems. They do not know there was a starting teacher. And they do not know that this teacher is not worthy of being called a teacher. That gives different results. This information should be on top of it [of the current adjustment procedure]. It is important for the school. ... Now it does not give a correct image of the school. (School 1)

The feedback was perceived as valid and reliable when the results were congruent with the findings of pupil monitoring systems, school tests, or intuition. When this was not the case, the results were seen as less valid, and low performance was more easily attributed to external factors, such as the difficulty of test items, atypical question methods, and incorrect results of the reference group (i.e., some schools were thought to have falsified their results by helping their pupils during the test). Others criticized the single-shot nature of the data gathering. A particular problem arose when a school was geographically distributed. Aggregation of data at the school level was of lesser value because the school's population, and sometimes also the school's policies, can differ between geographical locations. Finally, concerns were expressed when class organization or differentiation forms were very different from the approaches adopted in the reference group.

Differences in perceived user-friendliness of the SPFS. The nature of the overall feedback system influenced feedback use. Respondents complained about the large investment of time and effort during the data gathering process. Teachers and pupils perceived the tests as stressful. In addition, questionnaires directed to parents

required a considerable amount of time and a willingness to report private information. Furthermore, test times overlapped with other key assessment and evaluation periods in the school year. This explains why some teachers considered participation in the project as an extra burden on top of a heavy workload. This feeling was reinforced when the feedback was perceived as less relevant.

User-friendliness also refers to the tailoring of the school feedback. Some principals suggested adapting the report to the individual school setting. In the same line, satisfaction with the communication between user and the feedback system played a role. Moreover, the schools received the feedback at a rather unexpected moment, which made it difficult to include the new information in the policy-making cycle.

Support-related factors

The interviewees offered valuable information on user needs concerning feedback support and advised us about how to fulfill these needs. The results reveal that feedback use requires both policy-oriented and research-oriented skills. These are skills that must be developed (Visscher, 2002).

Differences in support needs. As mentioned above, most users reported not being able to interpret the information without extra support. Nevertheless, feedback support should go further than just assuring a correct interpretation. Almost all principals reported that they got stuck after their attempt at interpretation. They stated that they did not feel confident about their interpretative capacities and that they needed recommendations on how to proceed to the next phases in feedback use.

I can only hope my interpretations are correct. But definitely with the last report, it is so extensive that there is some – I will not say doubt, but fear – that it might be wrong. (School 6)

It is the same problem as with pupil monitoring systems. You can go to the teacher and say “these are your results. This child scores an E. Here you are.” That teacher will file the report and there it stops. (School 3)

The respondents asked for specific help dependent on whether they received positive or negative feedback results. Furthermore, they requested help in diagnosing the causes and circumstances that the results could be attributed to. Most respondents asked for concrete instructions for action. This suggests that consultation services could help to fulfill these additional needs.

Differences in support characteristics. When asked for ideas on how to organize support, some respondents requested a face-to-face introduction to the concepts and representations in the report. These sessions should be organized on site, but if that is not possible, regional meetings are acceptable.

Feedback support should be functional, offering intelligible, theoretical, and practical information. Principals expected the support to go beyond the interpretation phase and to empower schools to diagnose their results.

Concerning the interpretation, we try to manage it. But we do not know if we are doing it right. It would be interesting if the SPF project would come with the report to the schools and would explain the information in a team meeting with the teachers, with the whole team, to show us how to look at the results. “What’s the next step?” Because now

we only get the “sec” results and read them as such, as how they are printed. Even some reading advice is provided, the impulse to really do something with it is always lacking. (School 4)

Defining the role of external support services was a difficult issue. Some respondents claimed that schools have to take the lead in feedback use. This is in line with Earl and Fullan (2003), who claim that professional development will help strengthen personal confidence and self-efficacy in coping with complex feedback information. The respondents indicated a preference for internal support by counselors and via in-service training. External support from feedback suppliers should not interfere with these initiatives. They emphasized the demand-driven nature of support. This confirms the idea that external support must be tailored to the needs of individual schools. A sufficient level of goodness of fit is a requirement to achieve successful support (Nevo, 1995).

Principals also referred to school team members as a basis for support. Principals mostly got support from the special care coordinator or teacher. Often, these staff members were more experienced in interpreting statistical concepts and graphical representations. These staff members can play a role as complementary specialists. As they have a more flexible work schedule, they can allocate time to study feedback reports. This is not the case for teachers that have to work according to a prescheduled roster. Some school principals also wanted to protect team members against work overload, thus not involving them in feedback use activities. They might also have perceived these staff members as less important sources of support in feedback use.

Implications, limitations, and conclusion

The present study focuses on the perception of principals of school performance feedback and the actual use of feedback information. This study took place within the context of a larger project aiming to develop and implement a new school performance feedback system. This study also builds on an eclectic framework that integrates the literature on SPF use. This framework was the guiding structure for interviews with 16 principals from different primary school settings. Our results indicate that the elements presented in the theoretical framework reappear in the interviews. Figure 3 represents the integration of findings from the literature and our study.

The aim of this study was to illustrate and elaborate a framework of factors that influence school performance feedback use. Where previous studies have provided literature findings (Visscher, 2002; Visscher & Coe, 2003), perspectives of feedback suppliers (Schildkamp & Teddlie, 2008; Visscher & Coe, 2002), and quantitative methods of testing feedback use (Schildkamp et al., 2009), this study illustrates the influence of different variables on feedback use in a qualitative way.

From a theoretical perspective, our findings can help refine the description of feedback use. Whereas previous studies (e.g., Schildkamp, 2007; Visscher, 2002; Visscher & Coe, 2003) make a distinction between different kinds of information use (cf. instrumental, symbolic, and conceptual use; Rossi & Freeman, 2004; cf. strategic use; Visscher & Coe, 2003), an empirical investigation of the phases of feedback use has not been carried out. In this study, both were explored. Additional types of feedback use emerged from the data: a motivating and pupil-directed use. The interview data also show that different types of feedback use are related to one

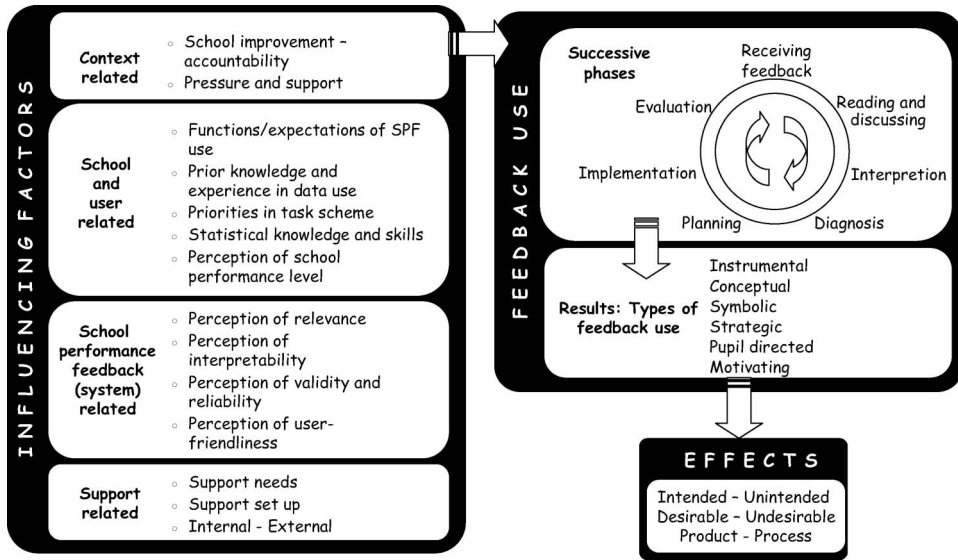


Figure 3. Integration of literature and research findings on SPF use.

another and occur simultaneously or successively. While a sequence of feedback phases can be discerned theoretically (Learning Point Associates, 2004), the process of feedback use is less systematic in practice. Our findings indicate that users can get stuck in the process of feedback use. A crucial challenge for future feedback use is to detect the difficulties in each phase and to offer appropriate support to systematize the process involved.

Our findings indicate that interpreting school feedback, making a diagnosis based on the results, discussing causes, and setting up actions based on feedback results is not a clear-cut process. The results reveal that feedback use requires both policy-oriented and research-oriented skills which must be developed by users (Visscher, 2002). Educational authorities should not neglect the importance of stimulating professional development and providing external support. Expectations about the positive impact that feedback use can have on school improvement will only be realized if extra support is available (Schildkamp & Teddlie, 2008; Sun et al., 2007).

To design appropriate support initiatives, a detailed analysis of the difficulties encountered when interpreting feedback reports must be conducted. For example, a recent study, which used both oral comprehension tests and IRT-calibrated online tests, illustrated the misconceptions that respondents reported during the interpretation of feedback reports. The results of that study contributed to the design of specific support initiatives (Verhaeghe, Verhaeghe, Vanhoof, & Valcke, 2009). Furthermore, experimental studies that manipulate the nature of external support can contribute to the design of a more sophisticated SPFS (e.g., Tymms, 1995) and the required support measures. In the design of SPFS, it is important to integrate the characteristics which appear to have a considerable influence on feedback use, such as relevance, interpretability, reliability, and validity. These characteristics are mediated by the perceptions of the feedback users. What is considered relevant by feedback developers, policy-makers, or researchers does not necessarily correspond with what the target group perceives as relevant. However, little is known about the effect of these differing perspectives in the context of school feedback use.

Moreover, one cannot expect schools to successfully implement innovations without making sufficient resources available (Davies & Rudd, 2001; Kimball, 2002). As school feedback use is not heavily promoted (Davies & Rudd, 2001), resources are limited. When we consider the work load of teachers and principals, our findings indicate that teachers will prioritize their classroom-related activities at the expense of school-level issues.

This study was conducted in Flanders, where there is no accountability culture or central examination system. It is not yet clear whether effective feedback use in such a context should only function within a school improvement perspective, as we found that feedback use was stimulated by an accountability orientation in terms of the inspection visits. It would be useful to examine the (in)direct influence of national and international authorities on feedback use (Creemers, 2006). Future research could focus on the relationship between a school improvement and an accountability orientation of educational authorities and key stakeholders (Vanhoof & Van Petegem, 2007) and on the balance between internal and external evaluations (Kyriakides & Campbell, 2004) influencing feedback use in schools.

The present study contains certain limitations. The validity of our findings is restricted to a specific educational context, with a particular school performance feedback system. However, the aim of this study was not to formulate generalizations but to explore and illustrate feedback use by its users. Another limitation is that a comprehensive framework is needed with an evidence-based set of influencing factors. Neither this study nor previous school performance feedback studies have attempted to meet this need. Furthermore, the link between school performance feedback use and the more general practice of data-driven decision-making remains unexplored. Despite the focus on accountability in the data-driven decision-making literature, common points of interest with SPF use can be further examined.

Acknowledgements

This research was supported by the Institute for the Promotion of Innovation by Science and Technology in Flanders (IWT), grant number SBO 50194. IWT is a government institute in Flanders stimulating and supporting innovation by providing financial support to research institutes.

Notes on contributors

Goedele Verhaeghe is a PhD candidate in education at the Department of Educational Studies, Ghent University, Belgium. Her dissertation research concerns the use of school performance feedback, and, more specifically, the influence of content and representation modes on feedback interpretation.

Jan Vanhoof is a postdoctoral researcher at the Institute of Education and Information Sciences of the University of Antwerp, Belgium. His interests are in educational quality assurance in general and in the conduct of high quality school self-evaluation in particular.

Martin Valcke is head of the Department of Educational Studies at the Ghent University, Belgium. His research field is the innovation of higher education and performance indicator studies. Detailed information about his publications can be found via <http://users.ugent.be/~mvalcke/CV/CVMVA.htm>

Peter Van Petegem is a full professor at the Institute of Education and Information Sciences of the University of Antwerp, Belgium, and is the head of the research group EduBRON (www.edubron.be). His research and writing interests include school effectiveness, educational quality assurance, school and teacher self-evaluation, and learning and instruction.

References

- Bosker, R.J., Branderhorst, E.M., & Visscher, A.J. (2007). Improving the utilisation of management information systems in secondary schools. *School Effectiveness and School Improvement, 18*, 451–467.
- Coe, R. (2002). Evidence on the role and impact of performance feedback in schools. In A.J. Visscher & R. Coe (Eds.), *School improvement through performance feedback* (pp. 3–26). Lisse, The Netherlands: Swets & Zeitlinger.
- Creemers, B.P.M. (2006). The importance and perspectives of international studies in educational effectiveness. *Educational Research and Evaluation, 12*, 499–511.
- Davies, D., & Rudd, P. (2001). *Evaluating school self-evaluation* (LGA Research Report No. 21). Slough, UK: National Foundation for Educational Research, Local Government Association.
- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London, UK: Routledge.
- Earl, L., & Fullan, M. (2003). Using data in leadership for learning. *Cambridge Journal of Education, 33*, 383–394.
- Fitz-Gibbon, C.T., & Tymms, P. (2002). Technical and ethical issues in indicator systems: Doing things right and doing wrong things. *Education Policy Analysis Archives, 10*(6), 68–82.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). London, UK: Cassell.
- Goldstein, H., & Myers, K. (1996). Freedom of information: Towards a code of ethics for performance indicators. *Research Intelligence, 57*, 12–16.
- Goldstein, H., & Spiegelhalter, D.J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 159*, 385–443.
- Hammond, P., & Yeshanew, T. (2007). The impact of feedback on school performance. *Educational Studies, 33*(2), 99–113.
- Hoy, W., & Miskel, C. (2001). *Educational administration: Theory, research and practice*. Boston, MA: McGraw-Hill.
- Kerr, K.A., Marsh, J.A., Ikemoio, G.S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education, 112*, 496–520.
- 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*, 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*, 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*, 23–36.
- Learning Point Associates. (2004). *Guide to using data in school improvement efforts: A compilation of knowledge from data retreats and data use at learning point associates*. Retrieved from <http://www.learningpt.org/pdfs/datause/guidebook.pdf>
- Liket, T.M.E. (1992). *Vrijheid & rekenschap: Zelfevaluatie en externe evaluatie in het voortgezet onderwijs* [Freedom and accountability: Self-evaluation and external evaluation in secondary education]. Amsterdam, The Netherlands: Meulenhoff Educatief.
- Lindlof, T.R., & Taylor, B.C. (2002). *Qualitative communication research methods* (2nd ed.). London, UK: Sage.
- Maso, I., & Smaling, A. (1998). *Kwalitatief onderzoek: Praktijk en theorie* [Qualitative research: Practice and theory]. Amsterdam, The Netherlands: Boom.
- Mason, J. (2002). *Qualitative researching* (2nd ed.). London, UK: Sage.
- Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Nevo, D. (1995). *School-based evaluation: A dialogue for school improvement*. Oxford, UK: Pergamon.
- Nevo, D. (2002). Dialogue evaluation: Combining internal and external evaluation. In D. Nevo (Ed.), *School-based evaluation: An international perspective* (pp. 3–16). Oxford, UK: Elsevier Science.

- Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. Burgess (Eds.), *Analysing qualitative data* (pp. 173–194). London, UK: Routledge.
- Rossi, P.H., & Freeman, H.E. (2004). *Evaluation: A systematic approach* (7th ed.). London, UK: Sage.
- Saunders, L. (2000). Understanding schools' use of "value added" data: The psychology and sociology of numbers. *Research Papers in Education, 15*, 241–258.
- Saunders, L., & Rudd, P. (1999, September). *Schools' use of "value added" data: A science in the service of an art?* Paper presented at the British Educational Research Association Conference, University of Sussex, Brighton, UK.
- Schagen, I. (2004, November). *Weighing the baby or fattening it: The use of data to inform school evaluation.* Paper presented at the NFER/ConfEd Annual Research Conference, London, UK.
- Schildkamp, K. (2007). *The utilisation of a self-evaluation instrument for primary education.* Unpublished doctoral dissertation, University of Twente, Enschede, The Netherlands.
- Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in The Netherlands: A comparison. *Educational Research and Evaluation, 14*, 255–282.
- Schildkamp, K., Visscher, A., & Luyten, H. (2009). The effects of the use of a school self-evaluation instrument. *School Effectiveness and School Improvement, 20*, 69–88.
- Silverman, D. (2005). *Doing qualitative research: A practical handbook* (2nd ed.). London, UK: Sage.
- Strauss, A.L., & Corbin, J. (2007). *Basics of qualitative research: Grounded theory procedures and techniques* (3rd ed.). Newbury Park, CA: Sage.
- Sun, H., Creemers, B.P.M., & De Jong, R. (2007). Contextual factors and effective school improvement. *School Effectiveness and School Improvement, 18*, 93–122.
- Tymms, P.B. (1995). Influencing educational practice through performance indicators. *School Effectiveness and School Improvement, 6*, 123–145.
- Van Aanholt, T., & Buis, T. (1990). *De school onder de loep* [The school under scrutiny]. Culemborg, The Netherlands: Educaboek.
- Van den Berg, R., & Ros, A. (1999). The permanent importance of the subjective reality of teachers during educational innovation: A concerns-based approach. *American Educational Research Journal, 36*, 879–906.
- Vanhoof, J. (2007). *Zelfevaluatie binnenstebuiten: Een onderzoek naar het verloop en de kwaliteit van zelfevaluaties in scholen* [Self-evaluation inside out: A study on the proceeding and quality of self-evaluations in schools]. Mechelen, Belgium: Wolters-Plantijn.
- Vanhoof, J., & Van Petegem, P. (2007). Matching internal and external evaluation in an era of accountability and school development: Lessons from a Flemish perspective. *Studies in Educational Evaluation, 33*, 101–119.
- Van Petegem, P., & Vanhoof, J. (2004). Feedback over schoolprestatie-indicatoren als strategisch instrument voor schoolontwikkeling [Feedback on school performance indicators as strategic instrument for school improvement]. *Pedagogische Studiën, 81*, 338–353.
- Van Petegem, P., & Vanhoof, J. (2007). Towards a model of effective school feedback: School heads' points of view. *Educational Research and Evaluation, 13*, 311–325.
- Verhaeghe, G., Verhaeghe, J.P., Vanhoof, J., & Valcke, M. (2009). *The value-added results of schools: How to represent school feedback information?* Manuscript submitted for publication.
- Visscher, A.J. (1996). The implications of how school staff handle information for the usage of school information systems. *International Journal of Educational Research, 25*, 323–334.
- 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. 41–71). Lisse, The Netherlands: Swets & Zeitlinger.
- Visscher, A.J., & Coe, R. (Eds.). (2002). *School improvement through performance feedback.* Lisse, The Netherlands: Swets & Zeitlinger.
- Visscher, A.J., & Coe, R. (2003). School performance feedback systems: Conceptualisation, analysis, and reflection. *School Effectiveness and School Improvement, 14*, 321–349.
- Williams, D., & Coles, L. (2007). Teachers' approaches to finding and using research evidence: An information literacy perspective. *Educational Research, 49*, 185–206.