

FEEDBACK OF PERFORMANCE INDICATORS: A TOOL FOR SCHOOL IMPROVEMENT? FLEMISH CASE STUDIES AS A STARTING POINT FOR CONSTRUCTING A MODEL FOR SCHOOL FEEDBACK

*Peter Van Petegem and Jan Vanhoof**

1. INTRODUCTION

In the last decade the development of (international) educational indicators has been extending in many ways: the number of countries involved, the aspects covered, the impact on educational policy, their attention in the media. More recent is the attention that is given to the feedback of indicators to individual schools. More and more stakeholders become convinced of the fact that a better use of the indicators could lead to powerful opportunities for individual schools to analyse and improve their educational quality (Van Petegem & Vanhoof, 2002). A very important role in this respect is reserved for models that provide individual schools with feedback.

This contribution firstly investigates whether there is a need for providing educational indicators and benchmarks resulting from international comparative research to individual schools. Next, two Flemish case studies that actually equip schools with individual feedback are presented. Based on reflection on the needs for feedback and the experiences in both the case studies a model for feedback is designed. In this third part of the paper some general and technical aspects of the feedback of educational indicators are described.

2. IS THERE A NEED FOR FEEDBACK?

There is no use in creating opportunities for feedback of quality indicators on and to individual schools, if there is no need for such information. As a consequence, the first question we need to address is whether schools themselves are interested in the indicators. The following elements indicate that -both from the government's and the schools' view- they actually are.

2.1. From the government's point of view

Schools are currently more autonomous than before. In the meanwhile (or as a result) they are expected to invest in internal quality control. A government that stimulates its schools to evaluate their own practices has to create the appropriate context to do so (Emin, 1995). Providing schools with relevant information concerning their own functioning is an important requisite to facilitate such evaluation. By providing indicators on individual schools the government primarily aims at informing schools in order to conduct a documented study of their own strengths and weaknesses. As such the

* The authors would like to express their appreciation for the efforts of the people involved in constructing the TIMSS and PISA school reports to provide the necessary data and for their clarifications and comments during the research project.

emphasis is not on judging the quality of individual schools from the government's perspective (cf. the inspectorate's task).

Another reason -from the government's point of view- that makes it desirable to provide feedback to individual schools is the motivation of schools towards the (international) gathering of quality indicators. Schools often participate in scientific research without noticing its immediate results. Despite the fact that the participation demands a considerable effort of schools they rarely see the direct benefits of it. Educational research that relies on the willingness of schools to participate is often confronted with a striking amount of schools that is reluctant or unwilling to cooperate. One of the main reasons to explain this reluctance is that principals and teachers are not convinced of the usefulness of the studies for their individual school. The feedback to and on individual schools creates interesting possibilities to alternate this situation.

2.2. From the schools' point of view

Schools that undertake a self-evaluation need appropriate information to do so. This information is needed in a form that schools can use to evaluate how well they are providing for the different needs of their pupils. On that basis they are equipped to evaluate their own performances and to take well-considered decisions about what actions and innovations to undertake. Having a realistic perception of its own efficiency and effectivity is an essential requisite to improve schools' quality. The confrontation of one's own performances with these of other (similar) schools -the mirror function of school feedback- creates powerful opportunities to stimulate quality development. The set of indicators enables schools to reflect and to discuss. The availability of benchmarks indicates whether they are performing relatively well or rather poor. In both cases schools are challenged to identify explanations, causes and -if necessary- solutions. Of course, the mere fact that schools receive individual feedback on their performances does not automatically imply that they will undertake the necessary actions. Feedback is a necessary step but it is not necessarily a sufficient step.

2.3. Two Flemish case studies - school reports based on TIMSS-R and PISA

Recently the Flemish Ministry of Education has adopted the policy that information deriving from comparative international studies should result in feedback to individual schools. Not that there is a legislative regulation from the Flemish Parliament or an official document that stimulates the composition of school reports. It seems crucially that the chairman and members of the Flemish steering committee of the TIMSS-R and PISA studies were convinced of the fact that the feedback of indicators is indeed an important possibility to enhance the impact of indicators. This interest for the feedback of indicators was the main impetus of the Flemish school reports. As such the existing feedback initiatives are rather the result of the personal interest of these members than of a formal policy decision. There is however another feature of the Flemish educational system that needs clarification. The TIMSS-R and the PISA studies are the only comparable indicators that are available in the Flemish educational system since it has no system of central examinations. This rather unique situation makes it comprehensible why Flanders has no previous experience in the feedback of nationally gathered information on indicators of individual schools.

In the following we will focus on two exemplars that actually provide feedback to schools on the basis of international research, namely the TIMSS-R and PISA school reports.

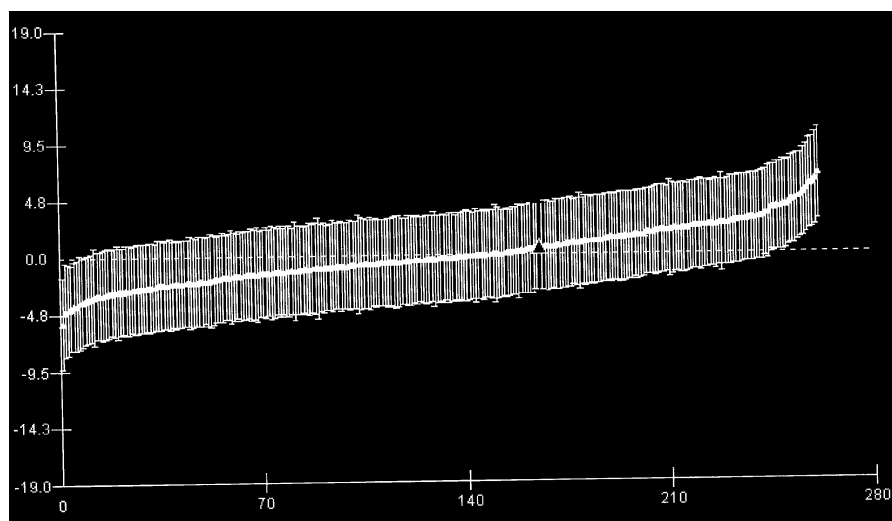
2.3.1 School reports based on TIMSS-R

The TIMSS-R project (Third International Mathematics and Science Study) focuses on pupils' knowledge of and attitudes towards mathematics and science in grade 8 (ISCED 2). The Flemish school reports discussed here are based on the results of the TIMSS-R (repeat) study (Van Damme & Van den Broek, 2000; MGv, 2002, pp. 24-34). TIMSS-R collected extensive information from pupils, teachers, and school principals about mathematics and science curricula, instruction, home contexts, but also school characteristics and policies. In Flanders researchers gathered some extra data in order to be able to explain the differences between schools and classes within Flanders. The extended data broaden the international study by including more classes and more pupils and by including additional variables (e.g. intelligence tests, the extent of problematic behaviour, and percentage of absence) and a questionnaire for the parents (including level of education, professional situation, country of origin).

The TIMSS-R school reports consist of three parts: an introduction, math and science results, and other relevant variables. The introduction focuses on some remarks that should be taken into account while interpreting the feedback. The other sections are more elaborated.

Concerning math and science results the questioned classes in a particular school are situated in relation to all the other classes (N=261) that participated in the TIMSS-R study. This comparison with other classes is provided on the ground of two types of information, namely on the basis of raw results and on the basis of adjusted results. In the latter comparison the intake characteristics of pupils are taken into account. Both the raw and adjusted class data are visualised using the following graph. The triangle (Δ) represents the mean score of the concerning class and the vertical line marks the confidence interval. The dotted line indicates the mean score of all Flemish classes.

Figure 1. Adjusted math results



The position of a class in such figure varies depending on whether raw data or adjusted data are used. Classes move over a certain number of places going from the raw data to the adjusted data. The school reports state that this means that pupils perform better or worse than would be expected given their input characteristics (Van Damme & Van den Broek, 2000). If the number of places a class moves over is negative this means that the class has a relative worse result when the scores are adjusted, if the number is positive the class has performed relatively better than would be expected

given the intake. On that basis the school reports aim to be able to judge schools to be more or less effective. The school report indicates how many places the individual classes in the particular school move when the results are adjusted. For example: Class a (2 Latin): number of places moved, -3.00 and Class b (2 Modern languages): number of places moved, +14.00. To interpret these data the TIMSS-R school report additionally presents a histogram that gives an overview of how many classes move up a particular number of places if the intake characteristics of pupils are taken into account. One graph focuses on math results, another on science results. This information aims at enabling schools to determine the relative quality of the results of the different classes that participated in the TIMSS-R study.

Several types of information concerning the math and science results of classes and their pupils may be obtained from figure 2. The school report provides the following clarification to this figure (Van Damme & Van den Broek, 2000, p.3).

Each pupil is represented by a little cross. The thick line is the most suitable connection between intelligence and math/science scores in the class. The dotted lines mark the 95% confidence interval. The long line demonstrates the overall connection between intelligence and math/science scores of all Flemish pupils in TIMSS-R. If the lower dotted line lies above the long line the class has a significant better math score than would be expected given the intelligence scores. If the long line is situated between the two dotted lines we cannot conclude with 95% certainty that the results of the particular class differ significantly from the overall results.

Figure 2. The correlation between intelligence and math/science scores

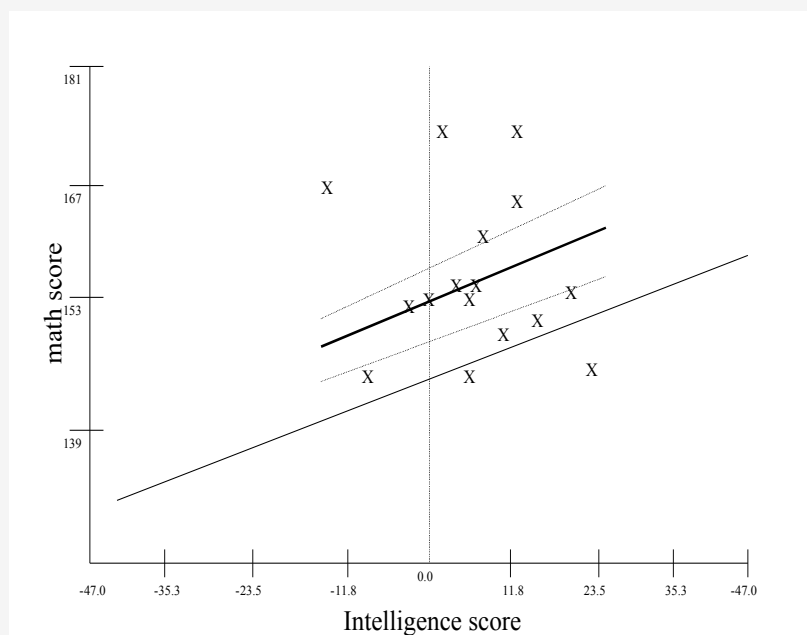


Figure 2 also provides information on the differential effectiveness of classes. This means that a class that realises relatively good results with low-intelligence pupils not necessary realises good results with high-intelligence pupils. This information can be deduced from the steepness of the lines. If the thick school line is less steep than the long line this means that in this particular class the pupils with a low intelligence score perform relatively better and the pupils with a high intelligence score relatively worse. If the thick school line is steeper the opposite is true. In the above example it appears that the class is not more or less differential effective than the mean Flemish class. The two lines are parallels.

Finally, the above figure also provides information on the input of classes. Both the length and the position of the

school line are relevant. The length is an indication for the heterogeneity of the intelligence scores of the pupils: a short line indicates all pupils of the group have similar scores, a long line means the class is composed of pupils with high and low intelligence scores. The position of the line –whether it lies mainly on the left or the right of the vertical 0-line informs about the actual intelligence scores of the pupils. If the line lies mainly on the left of the vertical dotted line, the class is composed of pupils with generally low intelligence scores. If it lies on the right the intelligence scores of pupils are higher than average.

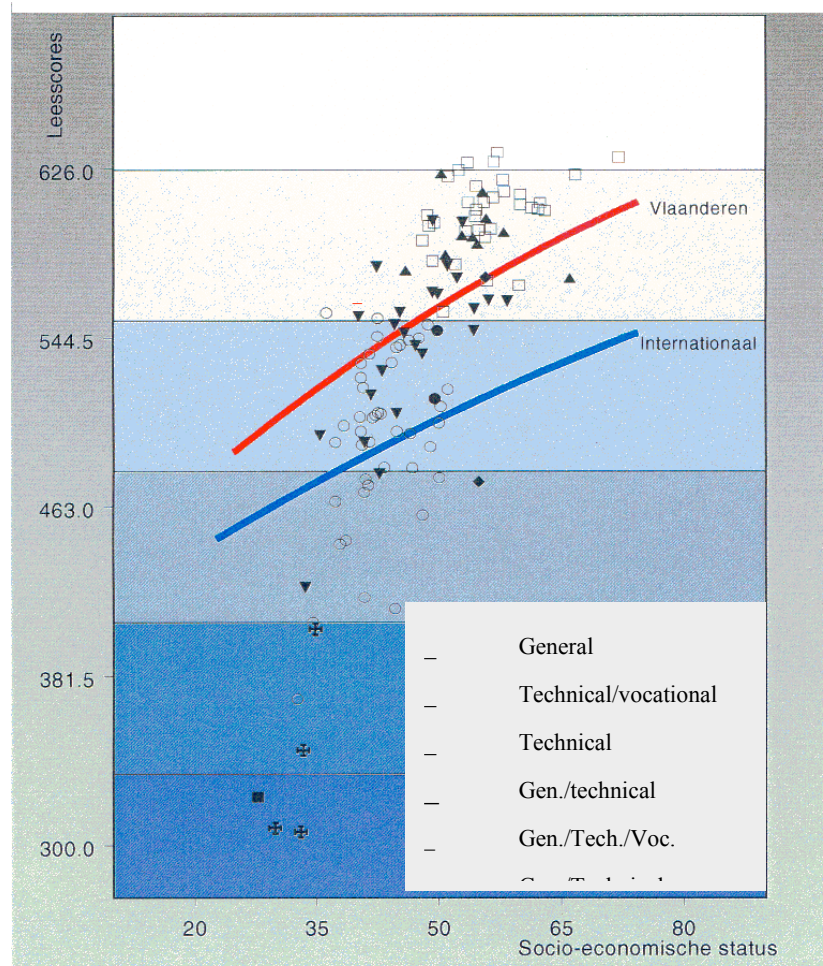
The third part of the TIMSS-R school report enables schools and classes to compare their results with these of all participating schools and classes. Information is provided at three levels: at the pupil level (index for the life comfort at the pupil's home, numeric and special intelligence score, education level of the parents and positive math/science attitudes), at the class level (teachers' view on the constructivistic organisation of the learning environment, obstruction of the instruction by the pupils and study orientation of the class group) and at the school level (frequency of problematic behaviour, the extent of problematic behaviour and percentage of absence). The information concerning these variables is presented to schools by reporting on (1) the results (mean score and standard deviation) of all participating classes and schools in general (2) on the results (mean score and standard deviation) of pupils, classes and the individual school in particular. In order to enable comparison the report provides graphs that specify how many classes or schools have a particular (mean) score. Based on this graph and the individual results schools can situate themselves in the broader group of all participating schools.

2.3.2. School reports based on PISA

PISA (The Programme for International Student Assessment) is an international survey on the knowledge and skills of 15-year-olds. Although the assessment domains are closely related to subjects learned at school, PISA concentrates on the value of the skills acquired, beyond the school gates. It assesses young people's capacity to use their knowledge and skills in order to meet real-life challenges, rather than merely looking at how well they have mastered a specific school curriculum. In the year 2000-study PISA assessed literacy in reading, mathematics and science. These data are however broadened towards more general outcomes and characteristics of learning. These include for instance: gender, family background (occupational status, family wealth, parental education, family structure, place of birth), and information on the learning environment and the organisation of schooling (school and classroom climate, learning outside school and resources invested in education). In Flanders 124 schools participated in the PISA-study (De Meyer *et al.*, 2002). In each school 35 random selected pupils took the assessments and were also asked to answer questionnaires about themselves and their schools. School principals were asked to give further information on school characteristics in another 30-minute questionnaire.

Each school that participated in the PISA study received information on the individual performances of its pupils. In the school report they received three drawings similar to the one in figure 3 (one for each literacy domain) (De Meyer *et al.*, 2002, p.21). Each school in the Flemish PISA study is represented by a symbol. The mark of the particular school is indicated in red and the performances of five similar schools are indicated in another colour. This way the researchers want to create an opportunity to compare the performance of a school with these of similar schools without indicating the name of these schools. The eight groups of similar schools are based on the (combination of) courses of study schools provide (for instance general education, technical education, general and vocational education, etc.).

Figure 3. The reading literacy performances of Flemish schools, in comparison with the Flemish and international gradients for reading



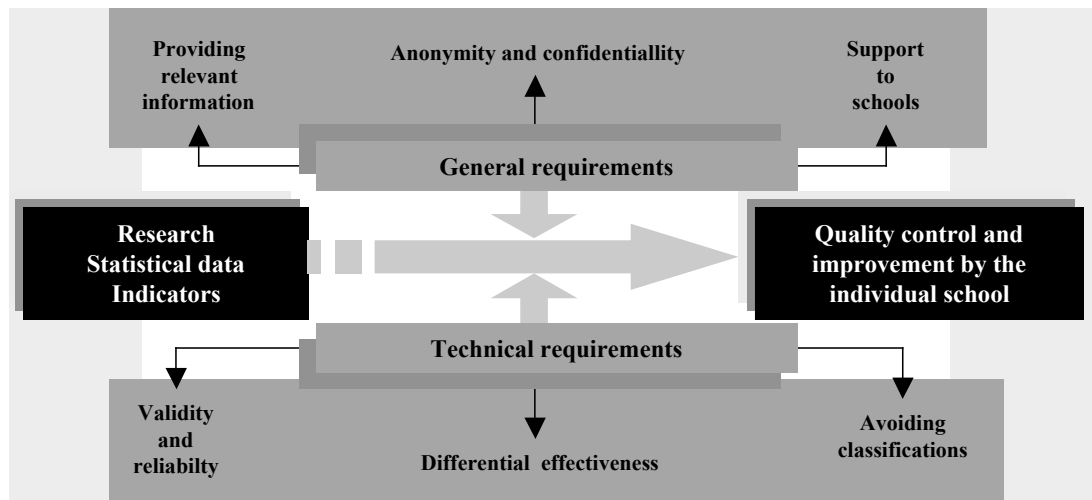
The symbol of the own school (indicated in colour) informs on two indicators. The height informs on the mean test score of all participating pupils in the school and the position in the breadth on the mean socio-economic status of the tested pupils. The different shades represent the different proficiency levels that are distinguished in literacy domain (six in reading and two in math and science). The graph also contains the Flemish (Vlaanderen) and international (internationaal) gradient of the specific literacy domain. As a consequence schools get an impression of their relative performance in a Flemish and international context.

3. CREATING MODELS FOR FEEDBACK

In the following paragraph we reflect on the above cases. This will result in a model for school feedback. The major aim of the model is to identify the conditions that have to be in place in order to make sure that a particular set of indicators may attribute to successful quality control and

improvement by the individual school. In order to successfully achieve that goal the model has to meet some general and some technical requirements (see figure 4).

Figure 4. General and technical requirements of the feedback model



General aspects of the feedback model encompass all modalities of the feedback system that are not related to technical aspects of the indicators. The most important elements relate to the content of the information, the anonymity and confidentiality of the feedback system and the support that schools receive in interpreting and using the information.

3.1. Providing relevant information - creating possibilities for self-evaluation

An obvious but complicated feature of feedback is content related. The feedback should be based on information that is considered to be relevant by the schools themselves. In order to determine which information meets this requirement we will present a general overview of guidelines that help to reveal which information is worthwhile to be incorporated in the feedback to schools. In general, the feedback of indicators has to create opportunities for schools to answer the two following questions: ‘how well are we doing?’ and ‘how well are we doing compared with similar schools?’ In order to answer these questions the following three features should apply.

a. Meeting the information needs of individual schools. The key to identify ‘how well are we doing?’ is having in place appropriate standards against which quality aspects can be measured, and sensible criteria to which those standards apply. This requires a balanced battery of available indicators and benchmarks. This enables schools to focus on these particular elements that they perceive to be most relevant. Schools have different priorities, different perceptions, different cultures but also different problems and as a consequence different solutions. This makes it advisable to give schools the opportunity to select to a certain extent the indicators that are of special interest. Feedback should invite schools to be selective and to make flexible use of the indicators. Does this imply that all data should be included in the school reports? Yes and no. Yes, all indicators can meet the information need of schools. No, only indicators that meet the criteria that will be presented in the following

should be considered. But yet, the challenge is to compose a balanced set of indicators that constitutes a useful and user-friendly tool. Finally, meeting the information needs of schools cannot be limited to informing schools on how well they are doing. Feedback should be broadened with indicators that provide information on how the school's quality can be optimized.

Although the test results in the TIMMS-R and PISA school reports might be an important aspect of a school's quality the school should put them along other information. Data on additional indicators are actually available in both the TIMMS-R and PISA dataset, but they are not included in the school reports. Especially the PISA school report is limited in this sense. There is a lot of information (indicators and benchmarks) that could be very interesting to schools while it is not included in the feedback.

b. Ensuring the possibility to compare and the comparability. The feedback should stimulate individual schools to question how better performing schools achieve what they do and what they can learn from those schools. This means that an indicator without reference to other schools or to good practices is not very useful to schools. In order to meet the information needs of schools the feedback has to enable schools to benchmark their position within the group of similar schools. In this regard it is very important that the feedback compares likewise. A mutual comparison of schools with a different profile is undesirable. The indicators used in the feedback should guarantee that all (or at least the most important) context and input aspects that influence school performances are taken into account. A good statistical analysis should be able to make a distinction between the effect of the school and the effect of pupils' background on test results and progress. Schools have to question their own performances on the basis of the differences caused by their educational quality and not on the basis of the differences due to the pupil's background (Yang *et al.*, 1999). Because of the misleading information from unadjusted quality indicators, the comparability of the available indicators is an essential feature of a good feedback model. After all, the comparison of non-adjusted output indicators will lead to false conclusions. Information that is based on poor comparisons will probably cause more damage than the absence of information.

Both the TIMMS-R and the PISA school report stress the importance of the possibility to compare school results with those of similar schools. The feedback of comparable indicators is however based at different procedures. The TIMMS-R indicator takes relevant intake characteristics of pupils into account. It is very important that the figure also contains the confidence intervals. This gives schools a clear overview of the real magnitude of the differences. The use of the indicator 'amount of places classes move over' can be misleading however. Statistical non-significant differences between classes can result in large changes in their relative positions. It would be more interesting to provide schools with information on (1) the mean score of their classes or school (including the 95 % confidence interval of this score) and (2) the number (or percentage) of schools that actually perform significantly better or worse. The PISA school report uses another method to compare schools. Instead of taking intake characteristics into account it compares individual schools with similar schools. This similarity is however solely based on the course of study the schools provide. Admitted, the social status is also indicated in the graphs but it remains impossible to determine whether schools differ significantly or not. School leaders and teacher have to be 'lucky' to find a

similar school to compare themselves with (concerning course of study and social status of pupils).

c. Ensuring the freshness of the data. Schools are not as much interested in the comparison of the results of their classes based on three or four-year-old data. They prefer more recent information, in order to sustain the value of indicators to question their actual performances. Schools want information that is perceived to be accurate and up to date. Educational practitioners and researchers know that there can be notable differences in the performances of different year groups (Bosker *et al.*, 1998). In order to come to correct interpretations of the indicators it is therefore important to hold the correct group of pupils in mind. Old data make this rather difficult.

The time span between the testing of pupils and the feedback of indicators is two and a half year for the TIMMS-R feedback and almost two years for the PISA feedback. This is a rather long but inevitable period. International research usually takes more time to gather, analyse and describe data than other research and there is a strict ultimatum to make the results public. But, while a two-year time span seems acceptable in the eyes of researchers, the question is whether schools themselves still experience the data as being accurate.

3.2. Ensuring anonymity and confidentiality

The school report should guarantee that schools cannot compare themselves with identifiable neighbourhood schools. Furthermore, providing feedback to individual schools has to take place in the absence of public involvement and pressure. This means that the information is confidential to the individual school and the (local) authority. Making the feedback to individual schools public might threaten the expected positive results and the internal quality improvement initiatives of schools (Vanhoof *et al.*, 2003). Feedback aims at inviting schools to optimise their functioning in a constructive manner. This requires a safe context for schools and teachers.

Both the TIMMS-R and the PISA school report consider the anonymity and confidentiality of paramount importance. The available information on other schools is presented in such a way that it is impossible to identify a particular school. Normally third parties cannot obtain individual school results.

3.3. Supporting schools in the use of indicators

Flemish schools have no experience in using quality indicators. Even rather simple statistics like the mean, standard deviation and percentile scores are not evident for a lot of practitioners. Neither principals nor teachers are automatically familiar with concepts such as confidence intervals, value added measurements, reliability, etc. They often do not understand the underlying statistical analyses and this threatens the correct interpretation of the indicators. Therefore explicit attention should be given to the support schools need to make maximal and correct use of the feedback. This support ought to focus on the correct interpretation of the indicators but also on the appropriate use of the feedback to evaluate their own performances. Interpreting indicators is just a first step towards successful self-evaluation. Schools need guidance in carrying out evaluations in their particular context, in selecting relevant performance indicators to evaluate specific issues and in using indicators within the development planning process (SOEID, 1997). Therefore, it appears to be advisable to provide schools also with step-by-step guides to evaluate the different aspects covered in the feedback.

These guides have to support schools by offering critical questions and activities that invite school teams to take a closer look at their performances.

The TIMMS-R and PISA school reports provide information to schools concerning the correct interpretation of the indicators. In both case studies the correct interpretation of the indicators is illuminated at a one or half day symposium and in the school report itself. This documentation seems clear and readily understandable for school leaders and teachers. As we mentioned before some presentations are however misleading. Schools are left with uncertainty about the different steps that could be undertaken in order to use the school reports as an input for reflecting on their own functioning.

3.4. Technical aspects of the feedback model

The model for successful school feedback also contains three rather technical requirements: indicators and their presentations have to ensure valid and reliable interpretations, should inform on the differential effectiveness of schools and should definitely avoid classifications of individual schools or classes.

3.5. Validity and reliability of indicators

In order to be informative, indicators have to be valid and reliable. Since one of the greatest dangers of indicators is the ease with which they can be misunderstood. Validity refers to the relationship between the measure and the inferences drawn from it. It is of primary importance that schools interpret the information in the school reports in line with what the data collection actually collected information upon. For instance, mere examination results of pupils allow valid interpretation on the examination results of pupils but not on the quality of instruction offered in schools. Reliability refers to the degree to which an indicator is free of random errors of measurement. That is, whether a student's score on a test is likely to change significantly from one administration to another; or whether statistics on a school's curricular offerings are likely to vary depending on who is collecting the data. If the requirements of validity and reliability cannot be met, the report should explicitly indicate the potential shortcomings of the available indicators.

Because of its specific purposes the interpretation of the PISA results for individual schools is somewhat problematic. The PISA test results are not only the outcome of schooling but also of learning processes outside the school and therefore they differ in some respect from the results of assessments focusing on the school curriculum (such as TIMMS-R). A school's PISA scores must be attributed to the cumulative impact of learning in schools but also in a range of other institutional and out-of-school settings. Moreover, the number of pupils per school is too small to ensure reliable measures at the school level. The PISA study does focus on the school curriculum and the (Flemish) data provide the opportunity to take into account some intake features of the school. As such it should be possible to make valid interpretations about the school's quality. It should be noted however that the PISA indicators are based on the performances of 13 year olds (second year of the first grade of secondary education). This means that pupils only attended the secondary school for one year and a half. This could be problematic because an important factor influencing the performances of pupils in secondary schools seems to be the primary school that was attended (Bosker, 1998).

3.6 Providing information on the differential effectiveness of schools

Within schools there can be significant differences in effectiveness according to different indicators (Goldstein, 1993; Fitz-Gibbon, 1997). Such differences refer to the so-called differential effectiveness of schools (Bosker, 1998) and may for instance be found between boys and girls, between different classes, different subjects, different prior attainment groups, and between different types of outcome (knowledge, skills, affective outcomes). These differences are not illuminated if one aggregates the available information in a general indicator of the entire school.

The TIMSS-R school report provides information on different subjects, groups of pupils and courses of study. As such it contains a balanced set of differential and aggregated data. The PISA school report provides differential information concerning the literacy results for reading, mathematics and science. Given the variables in the dataset it would be possible to provide schools with information (indicators and benchmarks) on their differential effectiveness towards gender differences, towards different courses of study and towards the different social status of pupils, at least if the number of pupils that participated in the study turns out to be sufficiently large.

3.7. Avoiding classifications of schools

Schools do differ, they have different results when compared on test results, even in the situation where pupil's intake characteristics are taken into account. An important feature of these differences relates to their magnitude. Only a minority of schools performs significantly better or worse than expected. The majority of schools will perform conform the expectations based on their intake. This means that the differences between schools are mostly rather small and that it is statistically worthless to present classifications of schools based on individual indicators (this is true for raw and adjusted indicators). Comparisons that present minor differences between schools as being notable (e.g. as in league tables) should always be regarded with great scepticism. Classifications create apparently big discrepancies between schools that actually have a quite similar performance. Because of the problems with classifications school reports should at least present confidence intervals in order to correctly interpret the differences between schools.

The TIMSS-R results reveal that there are differences between schools. These differences are not reported in mere classifications such as league tables. It is encouraging that confidence intervals are used. The PISA report does not present a classification of schools in a strict sense. However, the graphs with the literacy scores give no indication of the fact whether the differences in the performances of (similar) schools are statistically significant are not. As such the presentation might cause some problems concerning the interpretation too.

4. GENERAL CONCLUSION

As a conclusion we want to draw attention to some important lessons learned from the above reflections.

Informing schools themselves is the most important goal of the feedback on individual school data. Indicators and benchmarks aim at informing the school and its teachers in order to support them in carrying out self-evaluations. Each school should be in a position to make comparisons with other schools showing similar characteristics in order to assess its own performance and to decide what course to follow to improve standards and quality. Both from the government's and the school's

perspective this rationale for feedback is apparent. Another motive to provide feedback is the government's attempt to create and preserve the willingness of schools to participate in the time demanding research that is necessary to compose indicators and benchmarks. Looking at the Flemish situation it appears that –at this point- the feedback to schools is rather a poor comfort for their engagement in research than a user-friendly and elaborated tool for self-evaluation. This raises questions about the main goals of the Flemish government while providing such feedback to individual schools: are the school reports firstly intended to create opportunities for self-evaluation or are they –again at this point- just a sop for schools in order to thank them for their participation and to ensure their future cooperation.

In the light of the feedback initiatives at the national level in other countries (such as IPES in France, PANDA reports and Autumn Packages in England and 'how good are our results' in Scotland) the two Flemish case studies cannot be seen as what we could call exemplary feedback practices. Therefore they are not sufficiently elaborated. Of course, it is important to judge the usefulness of school reports in the light of the gathered information and the content of the dataset. If necessary data are not available it is impossible to provide schools with particular forms of feedback (remember that the Flemish government can not rely on central examinations in order to collect output indicators). Another aspect to be considered relates to the limited amount of time and costs that researchers are allowed to invest in the construction of the feedback.

Despite the limitations of the Flemish case studies we were able to make an inventory of some important features of a successful feedback model. The most important criteria relate to the content of the school reports. The feedback of indicators has to create opportunities for schools to question and analyse their own performances. In order to do so the use of benchmarks is promoted. Benchmarks are intended as a tool for schools by illustrating best practices and by informing individual schools on their relative performances compared to other similar schools. Such a comparison is nevertheless not simple. One has to compare likewise. This implies that a lot of attention should be given to the comparability of indicators. As we mentioned above: information that is based on poor comparisons will probably cause more damage than the absence of information. A lack of comparability is however not the only element that might threaten the successful use of indicators. Remember the emphasis we put on technical aspects such as the validity and reliability of indicators and the misleading information based on classifications of schools. Another paramount feature of feedback models ought to be the support given to individual schools. This support has to ensure the correct interpretation of the indicators and the appropriate use of the feedback to evaluate their own performances. In Flemish education –with few experience in the use of indicators- it is very important not to ignore the second part. But of course, there might be some cultural differences between educational systems concerning the importance of the different features the model should comply. Consider for instance the anonymity and confidentiality of feedback to schools. While this is probably a condition sine qua non in Flanders it might be that other education systems choose to make (some) indicators on individual schools available to other schools and the wider public.

REFERENCES

- Bosker, R., Lam, J.F., Luyten, H., Steen, R., & de Vos, H. (1998). *Het vergelijken van scholen [Comparing schools]*. Enschede, The Netherlands: Universiteit Twente.
- Darling-Hammond, L., & Ascher, C. (1991). *Creating accountability in big city schools*. New York: Columbia

- University, ERIC Clearinghouse on Urban Education, and National Center for Restructuring Education, Schools and Teaching.
- De Meyer, I., De Vos, H. & Van de Poele, L. (2002). *Worldwide Learning At Age 15 - First Results from PISA 2000*. Brussels: MVG.
- DfES, (2001). *Autumn Package*. London: DfES.
- Emin, J. (1995). *Des indicateurs de performance des lycées [Performance indicators on high schools]*. Paris: Direction de l'évaluation et de la prospective (DEP).
- Fitz-Gibbon, C. (1997). *The value added national project: Final report*. London: SCAA.
- Goldstein, H. (1993). Assessing group differences. *Oxford Review of Education*, 19, pp. 141-150.
- MEN, (Ministère de l'éducation nationale, de la recherche et de la technologie) (1999). *L'édition 1999 des indicateurs de performance des lycées [The 1999 Edition of Performance indicators on high schools]*. Communiqué de presse du 31 fev.1999. Paris: MEN.
- MEN, (Ministère de l'éducation nationale, de la recherche et de la technologie) (2000). *Les indicateurs de performance des lycées [Performance indicators on high schools]*. Communiqué de presse du 3 mars 2000. Paris: MEN.
- MGV (Ministerie van de Vlaamse Gemeenschap), (2002). *Vlaanderen in TIMSS [Flanders in TIMSS]*. Brussels: MGV.
- OECD, (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD.
- SOEID, (1997). *How good is our school? Taking a closer look at specialist services*. Edinburgh: SOEID.
- SOEID, (1998). *Taking a closer look at attainment in secondary schools in Raising standards – setting targets, Secondary Schools Support Pack*. Edinburgh: HMSO.
- The Scottish Office, (1997). *Raising Standards, Setting targets: Secondary Schools Support Pack*. Edinburgh: SOEID.
- The Scottish Office, (2002). *Setting Targets- Raising Standards in Schools*. Edinburgh: SOEID.
- Thomas, S. (1998). Value-added Measures of school effectiveness in the United Kingdom. *Prospects*, 28 (1), pp. 91-108.
- Thomas, S., Sammons, P. & Street, H. (1997). Value added approaches: Fairer ways of comparing schools. *SIN Research Matters*, nr. 7 (London, Institute of education, University of London).
- Van Damme, J. & Van den Broek, A. (2000). *Enige uitleg bij de schoolrapporten [Some clarifications by the school reports]*. Niet gepubliceerd manuscript. K.U.Leuven, Onderzoekscentrum voor Secundair en Hoger Onderwijs: Leuven.
- Vanhoof, J., Daems, F., Mahieu, P. & Van Petegem, P. (2003). *Naar publiek rapporteren over scholen? Lessen uit buitenlandse ervaringen [Towards publishing information on individual schools?]*. *School en Samenleving*, Afl. 2, februari 2003, pp. 5-25.
- Van Petegem P. & Vanhoof J. (2002). Feedback of indicators to schools. In: Scheerens J. [edit.]. *Benchmarking the quality of schools*. Enschede: Universiteit Twente.
- Yang, M., Goldstein, H., Rath, T en Hill, N. (1999). The use of assessment data for improvement purposes, *Oxford Review of education*, 25 (4), pp. 469 - 483.