

REPORT ON STAKEHOLDER PERCEPTIONS OF EMIS DATA

2010



basic education
Department:
Basic Education
REPUBLIC OF SOUTH AFRICA

ACKNOWLEDGEMENTS

The Department expresses its deep appreciation to Statistics South Africa for the technical support it has provided in the capturing and processing of data obtained in this survey.

ACRONYMS

ABET	Adult Basic Education and Training
ASS	Annual School Survey
ECD	Early Childhood Development
EMIS	Education Management Information System
FET	Further Education and Training
HEIs	Higher Education Institutions
SASQAF	South African Statistical Quality Assurance Framework

Contents

1.	Introduction	4
2.	Methodology	4
2.1	RATING CALCULATION	4
3.	Response	6
4.	Results	7
4.1	QUANTITATIVE FINDINGS	7
4.1.1	Methods used to access emis data	7
4.1.2	Frequency of use of emis data	8
4.1	QUALITATIVE FINDINGS	11
4.2.1	ACCURACY OF EMIS DATA	11
4.2.2	ACCESSIBILITY OF EMIS DATA	12
4.2.3	TIMELINESS	12
4.2.4	INTERPRETABILITY	13
4.2.5	RELEVANCE OF EMIS DATA	13
4.2.6	SUGGESTIONS ON IMPROVEMENT	14
4.2.7	GENERAL	14
5.	Conclusion	15

List of tables

Table 1:	Rating of timeliness for data related to ordinary schools	5
Table 2:	Respondents by market segment	7
Table 3:	Methods used to access EMIS data	7
Table 4:	Frequency of use of EMIS data in the past 12 months	8
Table 5:	Purpose for using EMIS data	8
Table 6:	Kinds of data used most often	9
Table 7:	Ratings (in percent) of EMIS outputs	10

List of figures

Figure 1:	Ratings (in percent) of EMIS outputs	11
-----------	--	----

REPORT ON STAKEHOLDER PERCEPTIONS OF EMIS DATA

1. INTRODUCTION

Over the past decade, the Department of Education (now reconstituted as the Departments of Basic Education and Higher Education and Training) has undertaken a number of annual surveys through its Education Management Information System (EMIS) to collect data about schools, Further Education and Training (FET) colleges, Adult Basic Education and Training (ABET) centers and Early Childhood Education (ECD) sites.

As is common practice world-wide, the Department saw fit to assess the quality of its EMIS data through a user satisfaction survey. The survey was also a response to the requirements of the South Africa Statistical Quality Assurance Framework (SASQAF), which obliges government departments to undertake user satisfaction surveys to obtain insights into perceptions of data quality by data users.

This report conveys the results of an on-line user satisfaction survey undertaken by the Department to assess the quality of its EMIS data. The on-line survey (**Appendix A**) was conducted during the period December 2009 and January 2010.

2. METHODOLOGY

An on-line questionnaire was e-mailed to 185 EMIS “users” who were identified through the EMIS outreach programme. These “users” comprised individuals as well as organisational representatives. Potential respondents were given a period of six weeks to respond to the survey.

Data capturing and processing was undertaken with the generous assistance of Statistics South Africa. Stats SA generated tables from completed survey forms and handed these to the Department for analysis.

2.1 Rating calculation

Question 9 required respondents to rate EMIS outputs on a scale from 1 to 4 (1 representing ‘very poor’, 4 representing ‘excellent’) in terms of relevance, accuracy, timeliness, interpretability, and ease of access. The following example illustrates the method that was used to transform actual responses into percentages:

Table 1: Rating of timeliness for data related to ordinary schools

Rating	Responses	Actual rating	Maximum rating
1	1	1	4
2	12	24	48
3	7	21	28
4	0	0	0
Total	20	46	80

Responses: shows the number of responses for each rating. In the above example, a total of 20 respondents provided a rating value for the timeliness of data related to ordinary schools. 7 respondents, for example, provided a rating value of 3.

Actual rating: the rating value multiplied by the number of responses (e.g., $3 \times 7 = 21$).

Maximum rating: this is what would have been the result if all the respondents selected the highest rating of 4 (i.e., the number of responses multiplied by 4) (e.g., $7 \times 4 = 28$).

The rating percentage is calculated by dividing the sum of the actual ratings (i.e., 46) by the sum of the maximum ratings (i.e., 80), and multiplying the result by 100:

$$46/80*100 = 57,5\%$$

Therefore, respondents rated the timeliness of data related to ordinary schools at 57,5%. This translates to an average score of 2,3 out of a total possible score of 4 (ie, the maximum rating).

2.2 Item response rate calculation

When interpreting the ratings reflected in the survey, one should keep in mind the item response rate. For each output (e.g. ordinary schools), a respondent was asked to provide a rating value (from 1 to 4) for five items (relevance, accuracy, timeliness, interpretability, and ease of access). If all 30 respondents had provided a rating value for all five ABET school items, for example, then 150 rating values in total (i.e. 30 questionnaires x 5 rating values) would have been available for analysis.

However, most respondents did not rate all outputs (which is expected, as a typical respondent might not be familiar with all outputs). Thus, some outputs received less rating values than others. The item response rate for each output was calculated by dividing the number of rating values actually obtained (in the case of ABET schools, respondents provided only 10 rating values across all 30 questionnaires) and dividing it by the total number of rating values that *could* have been selected for that output (i.e. 150).

$$10/150*100 = 6,7\%$$

Thus, out of 150 possible rating values that could have been received for ABET schools, respondents only provided 10 (6,7%). These 10 values were used to calculate the overall rating of 52, 5% for ABET schools.

3. RESPONSE

Of the 185 questionnaires e-mailed to potential respondents, only 30 responses were received, resulting in an overall response rate of 16%. Despite the low response rate, the findings of the survey are useful since many of the key and regular users of EMIS data known to the Department were among the respondents.

The following 22 organisations responded to the survey:

1. Centre for Education Policy Development (CEPD)
2. Chelsea Preparatory School
3. Department of Education, North West
4. Embassy of the Kingdom of the Netherlands
5. European Union Delegation
6. Human Sciences Research Council
7. North West University, Faculty of Education Sciences
8. Social Surveys
9. South African Institute for Distance Education (SAIDE)
10. South African Teacher's Union (SAOU)
11. Statistics South Africa
12. The Governing Body Foundation
13. The Presidency, Performance Monitoring and Evaluation Unit
14. University of South Africa (UNISA)
15. United Nations Children's Fund (UNICEF)
16. University of Cape Town, Children's Institute
17. University of Fort Hare
18. University of KwaZulu-Natal
19. University of Limpopo
20. University of Pretoria
21. University of the Witwatersrand (WITS), Education Policy Unit
22. University of Western Cape, FET Institute

Table 1 below indicates that, in the main, the survey was responded to largely by research organisations and universities (30% of total response, each), while international agencies comprised 10% of responses.

Table 2: Respondents by market segment

	N	%
National Government	4	13,3
Provincial Government	1	3,3
Parliament	0	0,0
Research organisation	9	30,0
Education Institution (e.g., University)	9	30,0
Community Based Organisation	0	0,0
International agency	3	10,0
Private Company	1	3,3
Media	0	0,0
General public	0	0,0
Constitutional institution/major public entry	0	0,0
Other	3	10,0
Total respondents (Questionnaires received)	30	100,0

Among the respondents were professors, senior researchers, senior lecturers, statisticians and Deans of Universities.

4. RESULTS

The results of the survey are divided into two main parts: a quantitative component and a qualitative one, as outlined below.

4.1 Quantitative findings

4.1.1 Methods used to access EMIS Data

Table 2 indicates the methods used by respondents to access EMIS data. More than one response was recorded for this question, hence the total number of responses (N) indicated in the table exceeds the number of respondents to the survey.

As shown by Table 2, EMIS publications and personal contact were the main ways in which respondents obtained EMIS information. Respondents do not appear to make much use of the internet to access EMIS information. This suggests the need for the Department to develop strategies to encourage and enhance greater use of the internet to improve its efficiency of information dissemination.

Table 3: Methods used to access EMIS data

	N	%
Telephone	12	13,2
Email	15	16,5
Personal contact	16	17,6
EMIS publications	18	19,8
Internet	15	16,5
Formal letter to the Department	14	15,4
Other	1	1,1
Total responses	91	100,0

Multi responses were allowed for this question, so the figures reflect the number of responses, not respondents

4.1.2 Frequency of use of EMIS data

Table 3 indicates the frequency of use of EMIS data by respondents in the past 12 months. According to Table 3, the majority of respondents (47%) indicated that they used EMIS data between 2 to 5 times in the past 12 months, while 30% indicated use about 11 times or more in the same period. Only 2% indicated to have not used EMIS data at all in the past 12 months.

It is apparent from the responses received, that EMIS data is used fairly often by some users, with close to one-third of respondents using EMIS data almost once a month, and close to a half using it a few times in a year. While it is encouraging to note fairly good use of EMIS data by respondents, it is equally important to strengthen the use of the data by stakeholders for the future.

Table 4: Frequency of use of EMIS data in the past 12 months

	N	%
None	2	6,7
Once	3	10,0
2 to 5 times	14	46,7
6 to 10 times	2	6,7
11 or more	9	30,0
Total respondents	30	100,0

Purpose for using EMIS data

Table 4 indicates the reasons why respondents use EMIS. According to Table 4, EMIS data is used mainly for research purposes (41% of respondents), followed by that of monitoring government policy (20% of respondents). A small, but significant proportion of respondents use EMIS data for planning (14%) and resource allocation (9%). The use of EMIS data for planing and resource allocation highlights the importance of ensuring that EMIS data is of good quality.

Table 5: Purpose for using EMIS data

	N	%
Research	23	41,1
Monitoring government policy	11	19,6
Planning	8	14,3
Policy formulation	7	12,5
Resource allocation	5	8,9
Other	2	3,6
Total responses	56	100,0

Multi responses were allowed for this question, so the figures reflect the number of responses, not respondents.

Kinds of data used

Table 5 indicates the kinds of data used most often by respondents. According to table 5, most respondents used information about schools, learners and teachers (20% in each case) in the past year, while very few respondents (3%) used EMIS data that reported on ABET. Respondents' use of data on Higher Education Institutions (HEIs) was relatively low (13%), suggesting:

- a) that other sources of information on HEIs are being used by respondents, or
- b) that the demand for information on HEIs is relatively lower than that for schools.

Table 6: Kinds of data used most often

	N	%
Information about schools	23	20,4
Information about teachers in schools	23	20,4
Information about learners in schools	22	19,5
Information about Higher Education institution	15	13,3
Information on ECD Institutions	10	8,8
Information about FET Colleges	7	6,2
Information on Special Need Education (Including special schools)	6	5,3
Other	4	3,5
Information about ABET centers	3	2,7
Total responses	113	100,0

Multi responses were allowed for the questions, so the figures reflect the number of responses, not respondents.

Rating of the quality of EMIS data

In line with SASQAF, the survey used five criteria to obtain information on how users rate the quality of EMIS data namely: relevance, accuracy, timeliness, interpretability and accessibility.

The overall response rate for this item was very low, with only 25% responses received for all outputs combined. This suggests that users tend to use data that applies to only specific levels of the education system. The section on ordinary schools had the highest response rate (66.7%) followed by HEIs with 35%.

Relevance reflects the degree to which statistical information meets the need of the user. Respondents rated the overall relevancy of EMIS data as excellent. Table 6 indicates that the relevance of HEI data was rated as being excellent (91%), followed by ordinary school data (82.5%). Data on ECD and ABET centres are rated as good (62.5%) by respondents. It is evident from Table 6 that respondents do not value the relevance of ECD and ABET data as much as they do with HEI and school-level data.

Accuracy of statistical information is the degree to which the output correctly measures the phenomenon it was designed to measure. The survey suggests that respondents do not have as much confidence in the accuracy of ECD and ABET data as they do with HEI and school level data. Overall, respondents rated the accuracy of all EMIS outputs, as good (63.3%). The

accuracy of higher education data was rated by respondents as excellent (77.5%), followed by data on special schools (75.0%). Data on ABET centres was rated as the least accurate (37.5%).

Timeliness is defined as the delay between the reference point to which the information pertains, and the date on which the information become available. Respondents did not rate the timeliness of EMIS data as highly as they did with other quality dimensions of the data. The overall rating on timeliness for all outputs was 56%. Timeliness of Special school data was rated good (66.7%), followed by Higher Education institution data (60%).

Interpretability of statistical information refers to the ease with which users understand statistical information by providing metadata. Users reported that EMIS data are easy to understand by rating interpretability good (70%). Interpretability of higher education institution data is rated excellent (80%) by users, followed by ordinary schools with 71%.

Accessibility of statistical information is defined as the ease with which users can obtain data from providers. Accessibility of all EMIS products is rated good (65%). Higher education institutions and Special Schools are rated excellent (75%) followed by public FET and Ordinary schools data with 65%.

Table 7: Ratings (in percent) of EMIS outputs

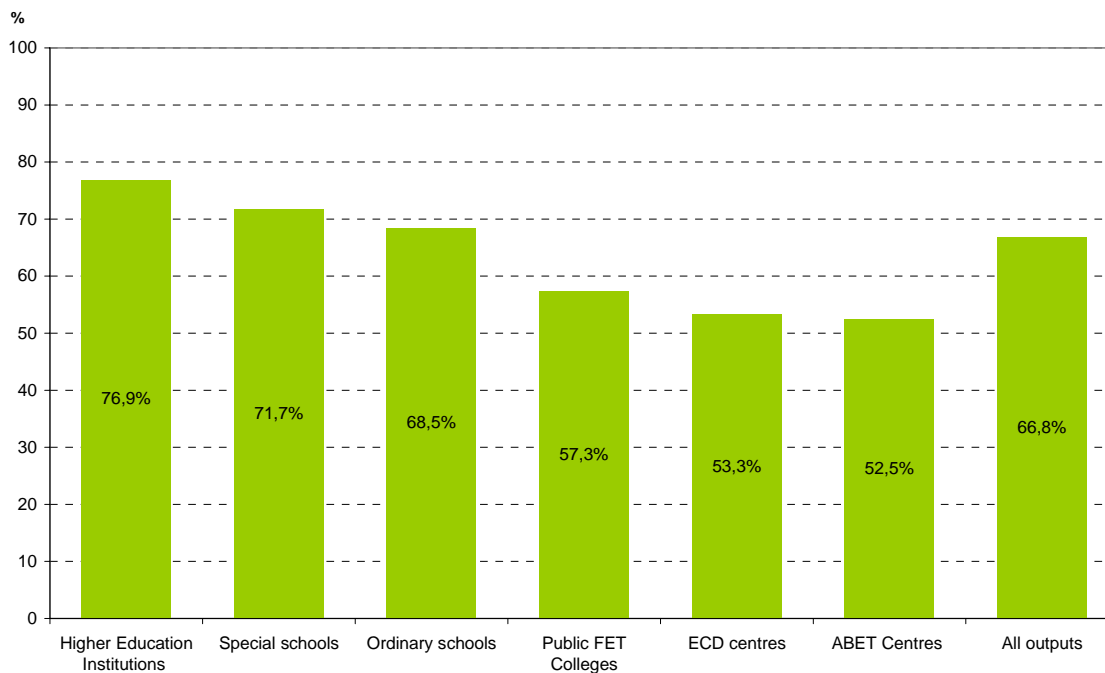
	%						
	Ordinary schools	ABET Centres	Special schools	Public FET Colleges	Higher Education Institutions	ECD centres	All outputs
Relevance	82,5	62,5	75,0	66,7	90,9	62,5	78,6
Accuracy	66,3	37,5	75,0	43,8	77,5	45,8	63,3
Timeliness	57,5	50,0	66,7	43,8	60,0	50,0	56,1
Interpretability	71,3	62,5	66,7	60,0	79,5	58,3	69,7
Ease of Access	65,0	50,0	75,0	65,0	75,0	50,0	65,4
Overall rating for output	68,5	52,5	71,7	57,3	76,9	53,3	66,8
Response rate (%)	66,7	6,7	10,0	16,0	35,3	20,0	25,8

Response rates for some of the EMIS outputs were quite low.

Figure 1 below represents overall rating of five SASQAF quality dimensions on EMIS outputs. Overall, the quality of EMIS outputs is perceived by respondents as being good (67%).

HEI information was rated by respondents as being excellent (77%) followed by special school information (72%). Given that school-level data (as opposed to data applicable to other levels of the system), is used most frequently by users (see Table 5), it is important that the quality of school-level data, even though regarded as being good by respondents, be improved.

Figure 1: Ratings (in percent) of EMIS outputs



4.1 Qualitative findings

In addition to requesting respondents to quantitatively rate the quality of EMIS data, the survey also requested respondents to provide comments on the quality of EMIS data in accordance with the dimensions of quality discussed above. The following are comments by respondents regarding their perceptions of the quality of EMIS data:

4.2.1 Accuracy of EMIS data

Although respondents rated the accuracy of EMIS data as being good in the quantitative aspect of the survey, they noted several concerns about the quality of the data in the “comments” section of the survey. Respondents pointed out that the database is not entirely clean and there are many missing data and duplicates.

Users stated that they experienced problems when merging two data sets. They mentioned that sometimes EMIS numbers in two datasets do not match. One user reported that the number of schools reported by provinces differs from the ones in the database.

Respondents also mentioned that some schools are listed in a category that they do not belong to (e.g. schools listed as being multi-grade when they do not offer multi-grade classes). While a number of respondents indicated that they trusted EMIS data, there was also concern about variability in the quality of data across provinces.

My sense is that data quality is variable. I trust most of the aggregated counts on learner numbers, educator numbers, age-grade figures and even absentee rates (more so than in the past). I also trust the school register of needs data (though this is now quite out-dated). Yet I'm hesitant to quote disaggregated figures from districts. Free State EMIS data is far more reliable and easy to access than Eastern Cape (for example).

Concern was also noted about the accuracy of information provided by schools. One respondent suggested that:

Provinces must verify data before submission to the DoBE - schools do not always complete ASS and Snap Surveys accurately.

4.2.2 Accessibility of EMIS data

Although respondents' overall quantitative rating of data access was good, some respondents highlighted problems they experienced in accessing the data as follows:

- I usually struggle to get someone to send the needed info and it takes time to receive it, often not clearly explained.
- Accessing data is often a frustration - especially where data is needed on a specific location or topic. Unless one has personal contacts in the provincial EMIS departments, accessing data can be difficult and time-consuming.
- Took many phone calls to get the data.

A number of respondents also indicated that they experienced difficulties when accessing EMIS data, because they found the Department's website difficult to use. One respondent had the following to say:

I find using the Department's website very frustrating. It is difficult to find what you are looking for and there is no search function.

4.2.3 Timeliness

Several users reported that sometimes the data they received, in particular, FETMIS data, was outdated. Reference was also made to the timeliness of the Department's School Realities publication.

- The School Realities information is excellent, but we would appreciate it, if it could be available earlier in the year instead of only in September.
- HEMIS and FET (particularly FET College data) data is not current so has very limited use.

4.2.4 Interpretability

A number of users suggested that it will be helpful to have easy access to metadata. Users also mentioned that clear definitions of terms should be provided to assist in data analysis:

- Easier access to metadata plus a regular publishing schedule will be a great help to the Department in improving these scores.

4.2.5 Relevance of EMIS Data

Users were generally satisfied with the relevance of information provided by EMIS. However, they provided useful suggestions on the kinds of information they believe is important for their use. Much of the information required by users can be obtained from the Annual School Survey (ASS) undertaken by EMIS, however the Department has not published this information as yet and is in the process of doing so. Below is a list of issues raised by users regarding the relevance of data:

Ordinary Schools

Learner Information: Users requested that disaggregated background characteristics of learners should be provided, as well accurate data on repeaters and dropouts. The following information was requested:

- Learner information by subject and grade: Grade 10-12.
- Drop out and repetition data at the national level. It is extremely time consuming to contact each province for the information and it takes months before the information is available.
- Disaggregated data of learners and teachers; data on orphans and vulnerable children disaggregated, retention rates for both girls/boys and MLA.
- Gender differentiated data by geographic area - where are girls in school / out-of-school?

School Information: Users requested that reliable data on no-fee schools should be provided and that schools should be classified as rural, urban or farm. In addition the following was requested:

- Reliable data on school fees (and fee revenue), collected by PEDs and integrated into the national EMIS.
- Number of classes per grade (i.e. not just number of learners) - to enable calculations of average class size.
- Multi-grade statistics

Staff Information: Users would like to see disaggregated information on staff including staff employed by SGBs and details of payments. Some users would like the EMIS form to reflect actual number of working hours of teachers. One respondent was particularly concerned about the absence of information on educator qualifications and commented as follows:

We utilized EMIS data for a research project for the DoE on teacher upgrading. The main area of interest was therefore on teacher qualifications - found in the last page of the annual school survey. It was found that in general the teacher information page was of secondary importance to the other components of the annual school survey as the focus is on the learner data. This meant that the PDEs focus on learner data and pay less attention to the consistency and quality of the teacher data and utilize fewer resources for its verification.

Another user had compared EMIS data with PERSAL data and found discrepancies between the two data sources with respect to REQV levels:

A comparison of the EMIS data with PERSAL data from one province for a similar time period revealed that there were many teachers who could not have provided the correct REQV level in their EMIS form. As a result, there was a limited amount that the information could be used.

ECD

Users reported that they would require information on the number of ECD centres and the number of qualified ECD practitioners by province and district.

FET

Some users noted that accurate and updated data on FET Colleges is not available. One respondent identified surveys that have already been undertaken in some provinces and proposed that the Department coordinate the process nationally:

ABET

Users requested that data on ABET centres, staff and learners as well as achievements be provided. They also suggested that EMIS should collect data on Adult Literacy.

4.2.6 Suggestions on Improvement

Respondents provided several suggestions on how to improve the quality of EMIS data. Suggestions include the need for a data cleaning system and comparing EMIS data with other surveys for the purposes of verification. One respondent even provided a report on a quality assessment of EMIS.

4.2.7 General

In general respondents were positive about the quality of EMIS data with respect to its access, accuracy, timeliness, interpretability and relevance. Despite some concerns expressed about the quality of EMIS data as enunciated above, many users were very impressed with the work of the Department and the usefulness of EMIS data, as indicated below by comments received:

- EMIS is an extremely impressive and useful resource, and we will continue to use it frequently.
- We admire the Department of Education's transparency in granting access to the data, and their open stance towards working with us to improve the status of Education statistics in the country.

5 CONCLUSION

The 2009 user satisfaction survey on EMIS data was the 1st such survey undertaken by the Department. The findings of the survey will no doubt be used by the Department to improve the quality of its survey processes.

Although the response rate to the survey was disappointing (30 respondents out of 185 known users), the Department was able to gain much insight into issues that were of concern to users of EMIS data from the qualitative part of the survey. It is hoped that an improved response will be received in the next survey.

This survey concludes that the overall quality of EMIS data can be considered to be good.

However there is patent variation in the quality of data pertaining to the various sectors of education (FET Colleges, Schools, HEIs and ABET Centres) as well as that between provinces.

It is apparent that data accessibility needs to be improved, mainly by ensuring that the Department's internet is more user-friendly.

Indisputable too, is the need for the Department to improve the accuracy of its data with the introduction of appropriate data cleaning systems.

On the whole though, users appear to be fairly satisfied with the quality of EMIS data, with some dimensions of quality being more satisfactory than others.