Curriculum, Evaluation and Management Centre

PUBLICATION NO.14

AN UP-AND-RUNNING INDICATOR SYSTEM

C.T. FITZ-GIBBON

1990

12 An Up-and-Running Indicator System

C. T. FITZ-GIBBON

A Performance Monitoring system will be described which evolved from a small research project started in a dozen comprehensive schools in 1983. By 1990 the system was serving most of the Local Education Authorities (LEAs) in northeast England and included more than 70 institutions (comprehensive schools, Colleges of Further Education, Sixth Form and Tertiary colleges) in seven LEAs.

The system monitors a discrete area of provision, namely A-level work. It is now known as the 'A-level Information System' (ALIS) but before LEAs became involved it was known as the COMBSE project, 'Confidential, Measurement-Based, Self-Evaluation'. Some results from the first year have been reported elsewhere (Fitz-Gibbon, 1985).

This chapter starts by looking at the end-product: what do schools, colleges and the LEA get by participating in the performance monitoring system? Subsequently the methods used to provide the information are considered, along with the costs of participating and the limitations of the system.

The A-level Information System

A school or college participating in the ALIS project receives, each academic year, three reports for each of eleven A-level subjects: Biology, Chemistry, Physics, Mathematics, Economics, History, Geography, General Studies, French, German and English Literature. The first report deals with examination results, the second with students' attitudes and the third with teaching processes and contexts.

Confidentiality is maintained by having all data reported under codenames chosen by each school or college. Thus the staff at each institution can see their own data in the context of all the data, without being able to identify specific other institutions. They are urged to choose codenames which are not transparent and to guard the codename as closely as their chargecard number.

Before considering the contents of the subject-by-subject reports we might note that the Head of each institution will probably pay attention to just a few indicators, leaving the Heads of Departments to consider the details in their curriculum areas. For the Head of each school or college, the 'first line' indicators (following the terminology adopted by Wilcox) will be those which answer the following questions:

1. Were the A-level examination results as good as should have been expected given the candidates with which each department was working?
2. Did students have a positive attitude to the subject they were studying at A-level?
3. Were students participating in a broad variety of extra-mural activities?
4. Did students have a positive attitude to the school or college?
5. Were students' aspirations for higher education in line with their abilities and achievements?

These first-line indicators do not provide a total picture of the institution but they do provide an efficient way to keep an eye on five major outcome indicators.

The first report: Examinations

The first report, on examinations, begins by introducing the project to new readers and giving general information about performance indicators. An explanation of the meaning of 'residuals' and 'controlling for intake characteristics' is provided, similar to that in Chapter 8. The data tables which follow the introduction are almost all 'league tables' showing data for each school in a rank ordered list. (There is other information in the table, such as levels of statistical significance but the rank-ordered list of averages is the primary information. The other information makes the tables look forbiddingly complex but serves two purposes: in addition to being of interest to people with a knowledge of statistics, it also serves to thoroughly discourage casual readers, e.g. journalists.)

Three data tables cover intake characteristics, showing the average per school (or college) for: O-level Grade Point Average, an ability test and the occupational level of the Head of Household. Thus Heads of Departments (HODs) can see how their intake compared with the intakes at other schools. They are encouraged by the accompanying text to find schools with comparable intakes and then, throughout the report, to compare their data with that from similar groups.

The next set of tables deals with the A-level results. One table simply shows the average A-grade obtained for each subject in each school, using an UCCA scale (A=5, B=4 etc.) This table is important in that it represents raw, unadjusted data which can be directly checked against the school's own record of the A-level results. Thus it serves as a validity check.

The next table reports the mean (average) residual for each school based on controlling for prior achievement (O-levels or GCSEs). For example, below is an extract from one report, showing residuals for eight schools for mathematics:

School code	*Residual*
NOCB	-0.5
WARD	-0.3
TRID	-0.2
WIRY	0.0
GEZI	0.2
BRID	0.3
HACE	0.5
QUIK	0.6

It is from this table that the Fair Performance Indicators for the subject are drawn. The text emphasises the need to avoid over-interpreting small differences and makes recommendations about the size of residuals which might be considered important, a figure which must vary from table to table depending upon sample sizes. Sample sizes cannot be shown in the tables because this could make the identification of schools possible.

The appendix to the report provides four more tables of residuals in which the following combinations of intake characteristics are taken into account: (i) the ability test (ii) the ability test and prior achievement (iii) the ability test and occupational status of the Head of Household and (iv) ability, prior achievement and occupational status.

Another appendix provides details of the regression equations along with descriptive statistics and intercorrelations. There are many numerate teachers well able to follow the simple regression procedures which have been used in the reports and this appendix is for their information.

Given the eleven reports, one for each of the examination subjects, Headteachers can examine the residual for each department. Knowing that positive residuals suggest the department obtained better results than 'expected' on the basis of the prior achievement of its intake, and a negative residual indicates worse results than 'expected', Heads find a summary such as the following helpful:

School code	*Residual for:* *ENGL.*	*FRENCH*	*HISTORY*	*GEOGR.*	*PHYSICS*	*MATHS*
BRID	-0.4	0.9	-0.8	-0.5	-0.2	0.3
NOCB	-0.8	-0.8	1.0	0.7	-0.3	-0.5
TRID	-1.0	-0.7	-0.2	0.1	-0.6	-0.2
QUIK	0.2	0.3	0.4	-0.2	0.5	0.6

These summary tables show that school departments vary considerably and no one school has the 'best' residual in every subject. Next year, also, the residuals will be different again. Thus the information in ALIS serves mainly to assist in monitoring school departments; it does not support the notion of setting schools in competition with each other.

The second report: Attitudes

This report explains how two attitude scales were produced by combining responses to several items (a technique explained by Hazelwood in this volume). Six items dealing with the students' attitudes to the school or college were combined to give a general 'Attitude to the institution' scale. Six items dealing with the extent to which students liked the A-level subject were combined to give an 'attitude to the subject' scale.

The tables report average attitudes per school in league tables as in the achievement reports. One item in the attitude-to-the-institution scale concerns whether or not the student would recommend others to take their A-levels at that school or college. As this is a directly interpretable item of considerable interest, the data are reported as raw percentages.

Participation rates in extra mural activities are reported simply as the mean numbers of activities per student in each group.

Another table presents correlations between attitudes to the school and to the subject and 13 other variables. The purpose of this table is to see if there were strong correlations between attitudes and levels of disadvantage, cognitive variables, measures of institutional size and type, levels of participation, and aspirations. (In general there were not which is why attitude scores were generally not adjusted in the way examination scores were.)

Finally the report considers destinations and aspirations. The percentage of students applying for universities, polytechnics, vocational training and jobs is reported as raw percentages so that schools and colleges can compare their percentages with those of similar schools or colleges. Such information might well be useful for careers guidance, particularly as patterns of entry to higher education change over the coming years. In addition to the raw percentages, the percentage likely to continue in higher education is reported *taking into account prior achievement and ability*. This table could alert schools which were under-encouraging students to apply for universities.

The third report: Processes

In this report there are no league tables. The report is designed to explore with staff the teaching and learning processes which appear to be related to getting good examination results (the residuals) and having students that liked the subject (the attitude to subject scale). This is the information which might be helpful in improving the delivery of education, that is in realising the major purpose of monitoring efforts: system improvement. The process reports also provide a framework in which the effects of such innovations as Supported Self Study can be directly assessed.

Class sizes, examination boards and time allocated per week are also reported, information which allows institutions to compare the context in which they operate with the general pattern. Since the major costs of

providing a course are covered by the data on class sizes and time-allocated institutions can compare their costs with the general pattern.

Data Collection: Costs and Methods

The collection and reporting of performance indicators requires test materials, questionnaires, travel. It uses up the time of people and computers. In other words, performance indicators cost money.

One major consideration is the cost of collecting the data from schools and colleges. Since students' attitudes are being assessed it is essential that confidentiality for each student is provided and is seen to be provided. We therefore collect the data ourselves. Data collectors from the university administer the questionnaire and give the ability test. Furthermore the data collectors use an audio-tape to standardise the administration. These precautions are seen as essential to protect the credibility of the data. They have the added advantage of making the data collection reasonably painless for schools: very little staff time is required.

Of course, the decisions on costs would be easier if one knew the benefits one was likely to be purchasing. Unfortunately it is only possible at the present time to guess as to the benefits of good monitoring systems. Until research has been conducted on the specific benefits of various kinds of performance monitoring, it would be wise for LEAs to ensure that the cost of collecting indicators is kept reasonably low. In particular, the costs in terms of the time of school personnel must be minimal, especially given the changing and sometimes tense relationships between LEAs and schools and the great pressures on schools at the present time with such changes as GCSE, the National Curriculum and Local Management of Schools.

One major way in which costs are kept down is by not collecting data from too many sources. For example, it might be interesting to obtain parents' perceptions but it would be a considerable further expense: another questionnaire, more organisation, a different analysis. Unless there is a clear need for data from sources other than the pupils, it makes sense to obtain all the data from the one source.

Other examples of cost considerations could be given. If, for instance, one took Coopers and Lybrand's appendix to heart (Coopers & Lybrand, 1988) one might have to measure staff demeanour. Problems of adequate sampling and the need for inter-rater reliability studies would lead inexorably to a methodologically sound solution: video cameras running at randomly

selected, unannounced times at various locations inside and outside the school. These could pick up unbiased samples of staff demeanour which could then be studied and rated by several raters. This would be expensive but it could be done. We are, however, not advocating it, and not only because of the expense!

Discussion: Features and Limitations of ALIS

ALIS is an information system with a limited focus. It is not a comprehensive monitoring of all aspects of provision. In particular, it deals with 'performance monitoring' as opposed to 'compliance monitoring' (Richards, 1988). It provides indicators of the quality of provision, not just indicators of quantity, and in doing so may help to guide improvements in quality. The system's features may be summarised as follows:

- monitors a discrete part of the general provision;
- the discrete part is one for which there is some consensus on goals;
- multiple outcomes are considered, not only examination results;
- the multiple outcome indicators have good behavioural implications: each student is valued equally (see Chapter 8);
- relevant input indicators are available to promote fair comparisons;
- process variables are collected, including student reports on classroom practice;
- data is pupil-level data specially collected under standardised conditions;
- confidentiality between schools is maintained by the use of codenames;
- the indicators are designed to identify strengths and weaknesses *within* schools, subject by subject, year by year, *not* to set schools in competition;
- the system provides feedback which generates ideas about ways to enhance teaching effectiveness;
- the system is designed to provide formative monitoring not summative judgments;
- the system developed over several years and is linked with on-going research

Like almost any mathematical model imposed on a complex system, the technique of regression analysis has limitations and inadequacies. Assumptions of linearity, homoscedasticity and that the predictor is without error can all be challenged. Furthermore validity studies which link quantitative indicators with richer qualitative investigations are urgently needed, as is the investigation of more sophisticated models (Goldstein, 1987; Raudenbush & Bryk, 1986). In the meantime the approach is fairer than the use of raw unadjusted grades.

Basically, the university provides what its personnel are best at providing: good quality data based on a research programme. The LEAs aim to provide the sensitive and effective management framework, a role well-suited to their advisers/inspectors. These advisers need some understanding of measurement, statistics and design concepts in order to interpret the data with authority and with all due caution: correlation is not causation.

Acknowledgements

The author is most grateful to the personnel in LEAs, schools and colleges for the numerous interesting meetings, discussions, and negotiations which have helped, over the years, to shape the project. Appreciation is also expressed to the Department of Education and Science for covering the costs of data collection for two years in the development of this project.

References

COOPERS AND LYBRAND 1988, *Local Management of Schools*. London: HMSO.

FITZ-GIBBON, C.T. 1985, A-level results in comprehensive schools: The Combse project, Year 1. *Oxford Review of Education* 11(1) 43-58.

GOLDSTEIN, H. 1987, *Multi-level Models in Educational and Social Research*. London: Griffin.

RAUDENBUSH, S. AND BRYK, A.S. 1986, A hierarchical model for studying school effects. *Sociology of Education* 59, 1–17.

RICHARDS, C.E. 1988, A typology of educational monitoring systems. *Educational Evaluation and Policy Analysis* 10(2), 106-16.