

ALIS IN BLUNDERLAND

C.T. FITZ-GIBBON

Journal of the Secondary Heads' Assoc.

1992

ALIS IN BLUNDERLAND

There is currently great political interest in – and little knowledge of – measures of school effectiveness. At the University of Newcastle upon Tyne, Carol Fitzgibbon has been working since 1983 on ALIS: an A-level performance indication system.

Quality assurance, performance indicators, value added . . . Do these current buzzwords represent an unwelcome intrusion of accountancy, business and industry into educational practice? Or are they ideas that headteachers would want to become involved with regardless of outside pressures? I suggest the latter. Why otherwise would almost every school and college in the North East be voluntarily participating in such a system?

The pressures for performance indicators are only just building up and yet schools and colleges have been interested and participating on a voluntary basis in the A-level Information System for some years now. Indeed ALIS is now spreading outside the North East, with participation from schools and colleges in Cambridge, Hampshire, Hertfordshire, Richmond, Staffordshire, Surrey and more. In many instances schools and colleges are actually buying into the system with their own LMS monies, rather than waiting for the LEA to join with them in the development of an effective monitoring system.

What is this 'effective monitoring system'? Underlying any system of monitoring, in education, will be the notion of the *effectiveness* of the school in achieving certain outcomes that are valued. The people who have been looking at school effectiveness are not in accountancy firms, nor are they in business or industry. They are in the educational research community. Educational researchers bring a concern for and a knowledge of schools to their work on performance

indicators. This leads to some stark differences between what has been recommended by accountants and what is recommended and adopted among educational researchers. Percentage pass rates, for example, figure heavily in the recommendations from accountancy firms but there are many things wrong with percentage pass rates: they take no account of differences between schools regarding the pupils that are entered for examinations; they encourage pushing out the weaker students whether

The indicator of effectiveness which is adopted by researchers counts each pupil equally

or not it is in the interest of the student or society to do so; they encourage teaching to the borderline, ignoring the student who is capable of an A (since that student is unlikely to fail and damage the percentage pass rate) even though the student capable of an A may need that grade very badly for his or her career. In contrast, the indicator of effectiveness which is adopted by researchers counts each pupil equally. It encourages bringing each pupil up to his or her potential. This encourages the kind of educational practice which is fair and desirable.

This particular feature of an indicator – the impact it has on behaviour – is the most important feature, and one which should be examined very carefully when any indicator is considered. The whole purpose of collecting information is to improve the system. If you collect percentage pass rates and set these up in league tables, the impact on the system could be to make it a less fair and a less effective system. Negative benefits could well ensue.

How do we make fair comparisons between schools? How do we construct

fair performance indicators and compare the value-added factor in each school? To begin with, we don't use the school as the unit. The data in the ALIS project shows very clearly that departments vary far more than schools in their effectiveness. The notion of a whole-school indicator should be vigorously resisted. It is not only in the ALIS data that the influence of the department on the examination results is found: a similar finding has been reported in Smith and Tomlinson's excellent book *The School Effect*.

So the question now is how do we make fair comparisons *among departments*? At A-level the matter is quite simple. The strongest predictor of performance in any A-level subject is the average GCSE score of the pupil. Not, note, biology GCSE to predict biology A-level, but the average of all the student's GCSE scores to predict the student's A-level grade. This is to be expected on theoretical grounds and turns out to be clearly the case in the data.

Suggestions that a sum-of-points score at GCSE should be used seem reasonable, but what of home background, or socio-economic status? First the socio-economic status of the catchment for a school is not a good

The correlation of home background measures with A-level performance is very small

indicator of the socio-economic status of pupils in any particular A-level group. Within a school, departments may differ substantially. In one of the most deprived schools, for example, we found a maths class in which all pupils came from professional backgrounds. We must look, therefore, at the intake to each department separately. Again, however, we must constantly be guided by the data and it

turns out that the correlation of home background measures with A-level performance is very small indeed. This is a considerable tribute to the education which is being provided. Pupils are being taught in a way that enables them to attain at A-level the grade you would expect them to attain on the basis of the GCSE

I wouldn't put any money on the SATs or NATs

scores, regardless of their home background. Of course it is a self-selecting minority which has stayed on to take A-levels. You might expect that the influence of home background would be greater on GCSE scores and that the influences are already incorporated in the GCSE scores to some extent but this is an endless argument back towards nature or nurture and it is simply not an argument which needs resolving. In an indicator system we are simply trying to say, 'What is a fair comparison for schools?' and comparisons are fair when the strong predictors have been taken into account.

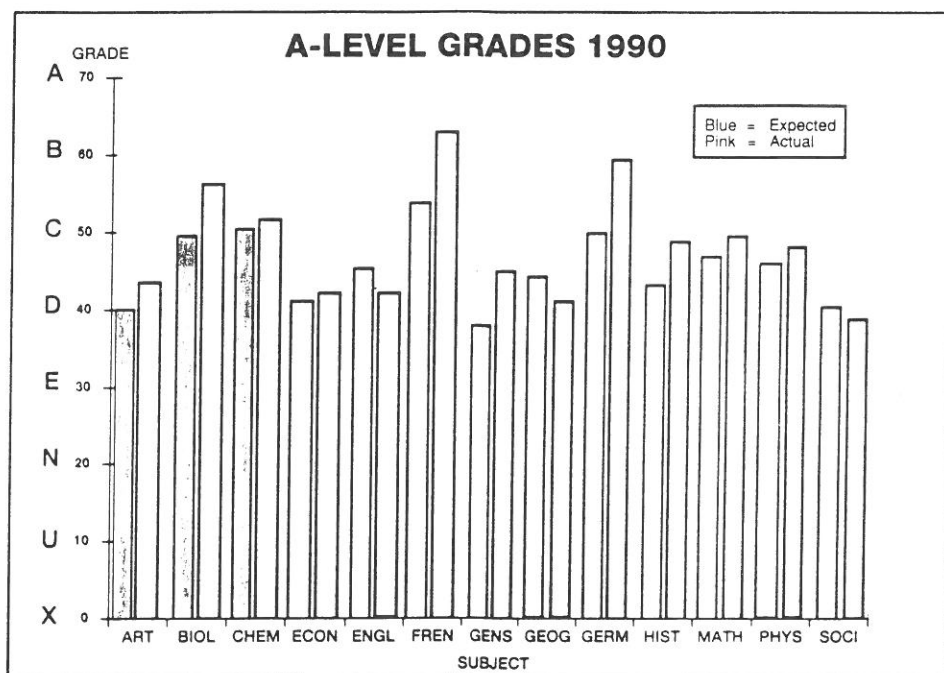
There are many problems in the measurement of home background, and it is therefore a relief that it seems not to be nearly as good a predictor as any cognitive measure. This then gives us the clue as to how to measure value added at year 11. A prior cognitive measure of some sort is needed. I wouldn't put any money on the SATs or NATs, or at least not so long as they are administered in school and the results have to be reported. This last statement may seem to cast some doubt on the tendencies of the teaching profession to act on Scout's Honour. When you construct a bridge, however, you have to construct it to stand in a gale, not just in a light breeze. When you set up a human system, it has to be incorruptible, otherwise nobody is quite sure whether it is being operated with 100% honesty. We should therefore challenge those who set up Scout's Honour systems. How are they going to find out if there is any cheating involved in the creation of performance indicators? If you can get away with cheating, you

probably should, just to draw attention to the deficiencies in the system.

The incorruptibility of the indicator is very important and, of course, in relying on A-level and GCSE examination results we are putting a lot of faith in the examination boards. Are they worthy of this confidence? The point is that while they might not be perfect, there is nothing better and they are in many ways very good.

There is one major flaw, however, in our otherwise excellent examination system. The flaw is that pupils' names and schools are on the papers as they are marked. Now those who mark papers may protest that they

comparison with any other method yet devised, examinations are nevertheless not the only outcome of schooling that we care about. If we set up an indicator system which looks only at examination results, the implication is that exams are the only factor of importance in schools. From the beginning therefore, the A-level Information System has assessed other aspects of the sixth-form experience. In a questionnaire administered by Newcastle University personnel, under carefully standardised examination-like conditions, students are asked in confidence to respond to questions about their sixth-form experience.



An ALIS printout for an individual school: average grades against ALIS predictions

ignore the name and they ignore the school, but Fahima Patel from Inner Town Comprehensive School may find her paper marked differently from that of Edward Smith-Carrington from a well-known independent school. Bias is unavoidable once a name and a school have been read. The name and school may often convey gender, ethnicity, social class and religion. It is a major flaw in the quality assurance procedures of the exam boards that they have not worked out a method of keeping the names and schools off papers as they are marked. Raffle tickets have managed to work out this problem – surely exam boards can do the same.

As splendid as examination boards are, and as reasonably incorruptible and fair as examination grades are in

Many questions were tried during years of development and some were discarded. (For example, it was surprising to find 'In bed' as the answer to 'Where did you do your homework last night?' This answer made the following question 'Were you alone?' look rather odd.) Over the years the questions were refined and developed into scales. Each year, the following information is reported back to schools for each department in the school:

- the attitude of the students to the subject;
- the attitude of the student to the institution (school or college);
- the aspiration level of the student;
- a quality-of-life indicator based on adding up the number of activities in which the students participate.

The aspirations for higher education

are an important indicator and they are reported in two ways as raw data and adjusted for GCSE scores. The adjusted data allows a school to answer the question: considering the prior achievement of our pupils are they aspiring for higher education to the extent that other similar pupils are in other schools?

The reports we give back to schools consist of league tables using code names chosen by the schools, so that each school sees all the data, but knows only its own code name. All this data makes for some fairly indigestible statistical reports. There is some advantage in this in that journalists show little interest in countless tables. However, within the schools these reports need to be backed up by INSET and this we provide.

There is a feature of the ALIS data which strengthens the sense that the data is there for collaborative investigation rather than for punitive surveillance or an indifferent accountability system. This feature is the presence of 'process variables'. Students respond on the questionnaire to questions about how often various classroom activities are used. The response scale is quite precise, not just 'rarely' or 'sometimes', but 'every

*Having students present
their work to the class
seemed to be a rather
under-used activity*

lesson', 'about once a week', 'once a fortnight', etc. This provides not necessarily totally accurate information about what goes on in classrooms (research is being undertaken on this question), but the students' perspective on the teaching methods. These 'process variables' are then correlated with exam outcomes and students' attitudes. It has been interesting to find, for example, that in chemistry classes where students reported being given essays once a fortnight or more frequently, the examination results were better than expected and the attitudes were more positive, not just in the data from one year, but in

three consecutive years, each year representing an entirely different set of students and possibly many changes in teaching personnel. This is just correlational data and does not tell us the cause of the relationship. But it is interesting information that

*The information . . . helps
the school to diagnose its
strengths and weaknesses*

raises questions about what essays mean, how they are set, how they are marked, what topics are used, and particularly, why essays set once a fortnight seem to be associated with the best value added and most positive attitudes of students. Another interesting finding in the process data was that having students present their work to the class seemed to be a rather under-used activity in many subjects.

In the year 11 indicator system which is under development, the concerns for outcomes include aspects such as whether pupils feel safe in school and, if they don't, where in school they feel unsafe: the corridors, the toilets, the buses? This work grew out of a TVEI-sponsored project on 'diagnostic data' and this aptly chosen title conveys the basic principle that the information which is provided in these indicator systems helps the school to diagnose its strengths and its weaknesses. It helps the school to keep a watchful eye on the trend in indicators from year to year. It helps the school to know where help is needed and where praise should be given. In this spirit of collaboration, no one is saying there is a single correct method to teach; nobody is claiming to know 'good practice', while not divulging the basis of this knowledge; no one is bullying teachers as to whether or not they use dictated notes. An eye is being kept on the outcome. If the outcome is good everyone can feel confident that somehow they are getting it right.

Because these indicator systems are being run as a research and development project with Newcastle University, the problems in the system

and the limits on the interpretation of the data are emphasised in the in-service workshops. There is no pretence that these are easy techniques to get right. Furthermore, the project is run very economically and gains considerably from research projects undertaken by students at no further costs to those participating in the indicator systems. A further aspect of the cost effectiveness of these systems is that we provide for each department a pupil-by-pupil measure of value added. If some pupils are working through supported self-study or as part of a consortium and studying elsewhere, or if some students came down with long bouts of glandular fever, all these particular instances can be taken into account by simply grouping the value-added measures for each student in an appropriate fashion. Thus the effect of supported self-study can be examined by looking at the value-added measures for those students who were or were not

*The growth of monitoring
systems may well be the
major advance in
educational practice in
the 1990s*

studying a particular subject using supported self-study methods. In this way, monitoring systems provide a quantitative framework of carefully collected, high-quality information. In this framework evaluation activities can be effectively undertaken.

The growth of monitoring systems may well be the major advance in educational practice in the 1990s. Linked with appropriate qualitative on-site methods and given the continuing use of careful outcome measures such as examinations, it can be expected that monitoring will contribute significantly to our being able to show how effective our departments are and to see off the latest fashions which are thrust upon education without any justification or evidence to back up the exhortations. Monitoring empowers the profession. ■