



Curriculum, Evaluation and Management Centre

PUBLICATION NO.07

QUALITY ASSURANCE SYSTEMS
IN EDUCATION

C.T. FITZ-GIBBON

&

P. B. TYMMS

Quality and It's Application
1993

QUALITY ASSURANCE SYSTEMS IN EDUCATION

C T Fitz-Gibbon and P B Tymms

School of Education, University of Newcastle upon Tyne
St Thomas Street, Newcastle upon Tyne, NE1 7RU

The A Level Information System (ALIS) was developed in Newcastle over ten years ago and now covers much of the UK. Its origins and the ideas behind its development are discussed. This quality assurance system is contrasted with the government initiative on national testing which the teaching profession has strongly rejected. Also the very different philosophy behind the new arrangements for inspection are juxtaposed with the spirit of research in the workplace which ALIS hopes to foster.

The concept of 'quality' in education must be rather different from the concept of quality in more simple production processes, such as in manufacturing the ubiquitous 'widgets'. The output for the educational process is a citizen and since there is a universal right to education, the characteristics of the input to a school — primarily the pupils — cannot always be selected. Since the input has an overwhelming impact on the output, absolute judgements about the output do not provide indicators of the quality of the process. Hence the interest currently being expressed in measures of relative progress made by pupils, often referred to as the 'value added', by the school. To obtain 'value added' measures for various stages of education, measures of both input and output are needed, such as achievement at age 14 and achievement again at age 16. Recently the design of measures of achievement has gone badly wrong in the UK, as witnessed by the fact that the tests which were to be given to all 14 year olds have been almost completely boycotted by teachers.

"At 1pm last Monday, after months of trying to head off a teachers' boycott of the first national tests for 14-year-olds in English, mathematics and science, John Patten, the beleaguered education secretary, was finally forced to admit defeat. Out of 600, 000 pupils only a few thousand pupils in a handful of schools put pen to paper. The tests were a flop." (Sunday Times, 13-6-93, p.11)

The government planned to have tests at ages 7, 11, 14 and 16 conducted in schools and reported in 'League Tables' (rank ordered lists) in the Press.

This breakdown in the plans, due to the boycott by teachers, gives the public the impression that teachers are resisting procedures which are putatively being introduced to produce "quality", to improve standards, to open up schools to parents. Thus the dismay in the Sunday Times (a pro-government newspaper):

"How could it be that such a simple reform — one of the 'big ideas' of the Thatcher era — was so comprehensively undermined? If successive Tory governments were able to defeat the miners, privatise public utilities and curb the excesses of left-wing local authorities, why had they failed to ensure children are taught to a national standard and tested on whether they can read and write?"

(13 June 1993, p11)

How indeed? And how does this supposed opposition of the teaching profession to accountability square with our own quite contrary experience at Newcastle where we have run a value-added system ("ALIS", to be explained later), based on examination results, for over a decade. The information we provide forms a statistical basis for quality assurance and is a system so welcomed by teachers that, from small beginnings with a dozen schools, it has now grown to over 368 schools and colleges from around the UK. They are participating voluntarily in the system, in many instances actually choosing to spend their own limited disposable income in order to participate. For explanations we must consider the whole system involved in the collection of such data.

students largely takes place externally, independently and 'blind'. The system is, therefore, seen as fair and incorruptible. Teachers and students know the syllabus which is to be covered and have had sight of previous examination papers. The examinations are taken under standard conditions throughout the country at the same time, on the same day, and the papers are marked not by the teachers but by paid examiners who do not know the candidates. The same cannot be said for some other educational assessment systems and monitoring exercises in those areas must be correspondingly hampered.

In the next section of this paper, we will consider how to set up a system of performance indicators which can provide the first stage in a quality assurance system. The performance indicators will provide the statistical data which is at the heart of many quality assurance systems; data rather than opinion. In considering each design stage, the way in which that stage was dealt with in the A Level Information System will be described.

Adopting an Approach:

Purposes and methods of working

The purpose of setting up an indicator system is clearly to assure quality and to, as Deming [3] says, strive for constant improvement. We therefore have to ask what kind of methods of working will produce the desired outcome in the entire system. What kind of use of statistics will improve a complex system of interacting humans? As we see it, the way to ensure quality is to pursue the approach and method of working adopted in 1983 when the system was first put into operation. At that time, it was called the 'Confidential Measurement Based Self Evaluation Project' (Fitz-Gibbon, 1985) [4]. **Confidential** because participating schools saw all the data but under code names which only they themselves knew; code names such as Eton and Colditz chosen by the schools themselves. **Measurement based** in that no presumption to be able to judge what styles of teaching or of organisation in a school would assure the best outcomes. We would simply measure and relate processes to outcomes statistically, rather than presume to judge the process itself. The presumption of knowledge is a great barrier to the acquisition of information. **Self evaluation** referred to our acknowledgement that the interpretation of the data would have to be made in each school itself. We could give a high quality framework of data but could not know all the particular events and circumstances in every class, every year.

We could not know, for example, that half a dozen pupils were absent with glandular fever from an A Level Economics course one year, and that this most likely affected their achievement. The statistical monitoring could not be operating at that level of detail. In summary then, in the ALIS project, the purpose is to collect data about how the system is functioning and feed it back into schools under code names, confidentially, for their own investigation and interpretation of the data.

Rationale for the Choice of Performance Indicators

There is a great danger of regarding any piece of information that drops off a computer as a potential performance indicator. Such an approach leads to drowning in data and an unproductive accumulation of uninterpreted information. Before data is collected, tentative models of how the system is working must be adopted. We approached this problem with the following rationale. Firstly we asked of ourselves, and of the schools and colleges with whom we work, what are the outcomes that we care about enough to attempt to measure them? At A level these were:

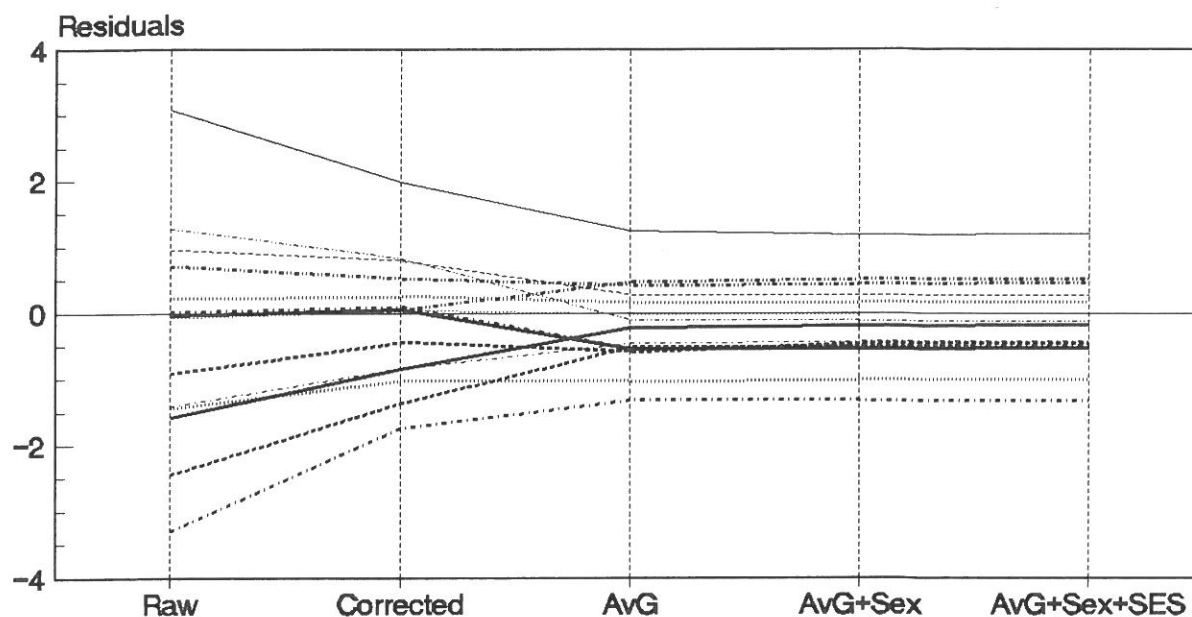
- (A level grades
- (Attitude to the Institution
- (Attitude to subjects studied
- (Quality of life in the institution
- (Aspirational levels

The next step was to consider what features of the situation, which are out of the control of the school, will affect those outcomes and should, therefore, be taken into account? In other words, this is the search for covariates, the influence of which must be measured in order that the outcome can be fairly interpreted in a comparative fashion across schools. An example in relation to A level grades is given in the next section.

The ALIS system takes the department as the appropriate unit of analysis as far as the reporting of indicators is concerned. This can be supported by two arguments. This first concerns the nature of educational structures where decisions about how to teach and even what to teach are taken at the departmental level. The other rests on the data and that will be taken up in the next section. The quality control charts then should map the effectiveness of departments, establish reasonable bounds for common cause variation, and take action only if there is reason to believe there is an exceptionally effective or exceptionally ineffective department, i.e. if a

Diagram 1

Biology A level 1992



strategy to make it look as though they were working with less able pupils and, therefore, the value added would be more commendable.) In addition to being incorruptible, valid, credible and non-corrupting, it is important that data collection is as economical as possible, and as undemanding of the staff as possible. Teachers are busy teaching and do not have time to devote to data collection.

The Data Analysis

It has already been mentioned that the data analysis must be such that important links in the system are modelled in the data. Quality occurs in the classroom and the data must be fed back to the unit which manages the classroom, this is generally the department. So, the unit of aggregation for the production of reports in the ALIS system is the school department as, for example, the English Department, the French Department, the Maths Department, the Physics Department.

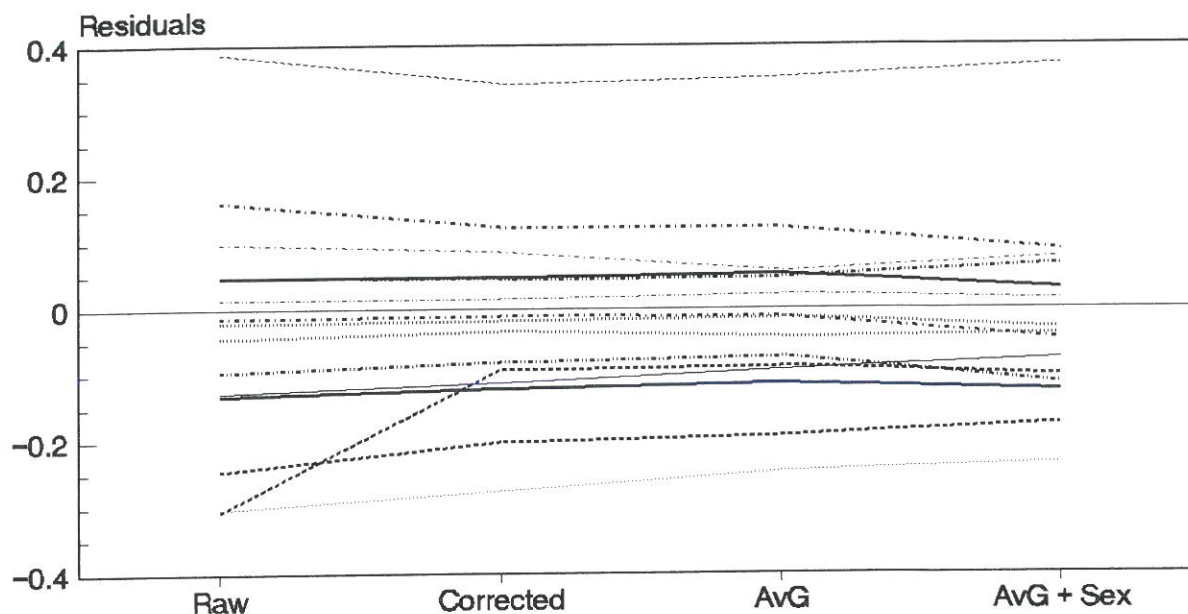
There are decisions to be made on how to analyse the data, for the purposes of feedback, taking account of the need for the data to be interpreted in the schools. In other words, there is an issue of some tension between complexity and thoroughness versus transparency to the user. Our solution is to go for both. One table provided to schools rests on a simple regression of A Level grade against average GCSE grade, a regression which is easily explained. Other

tables rest on multiple regression, and still other analyses are conducted using hierarchical linear models such as the ML3 programme from Harvey Goldstein and his colleagues at the London Institute of Education. An example of one of the tables is shown in Diagram 2.

Reports are prepared with many such tables showing intake, process and outcome variables for examinations, attitudes, aspirations and for the relationships between processes and the major outcomes. Each department receives a report of its own. It might be cheaper to put all the reports into one tome, but our commitment to feeding information back to the unit of management, back to the unit of responsibility has led us to produce separate reports for each department. Staff also receive pupil by pupil lists showing the average GCSE for that pupil, the predicted A Level grade in each subject, the obtained A Level grade and the difference between these two, that is the 'residual', which in modern parlance has come to be called the 'value added'. The regression is calculated anew on each year's A Level data relating it to achievement two years earlier in GCSE examinations.

Diagram 3

BTEC National Diploma



qualifications where the final grade is often heavily dependent on the course providers. Diagram 3 shows the results of constructing multi-level models of BTEC National Diploma qualifications. Very little difference can be noticed when the average GCSE and gender were added to the models; an indication, perhaps, of questionable measures. On the other hand one might argue that vocational qualifications should not have as much relationship to prior academic measures as A levels do. Nevertheless correlations as low as .2 are surprising. The data also indicated that membership of courses was associated with an unexpectedly high 40% of the variation in course qualifications — more usual figures in educational studies would be between 5 and 20% — and this strengthens the view that these internally assessed vocational qualifications are excessively dependent on the grading of the lecturers or teachers who run the courses.

The monitoring of educational provision is a complex process and the construction of monitoring systems must reflect that complexity whilst drawing strength from simplifications which can result from a careful consideration of the data. But the data itself must be quality assured — simply to use what is available may not be adequate and could lead to erroneous conclusions.

CONCLUSION

The accountability system based on consulting students in confidence, collecting examination data and analysing it fairly has proved very acceptable to schools and colleges. The system has doubled in each of the last three years, and the Secondary Heads Association invited us to provide the examination indicators for their members leading to another system called the SHA Value Added Project with over 350 schools in that, including many from the independent sector. Tymms (1992) [13] has shown few substantial differences between the value added measures and the type of school. The idea that creating new kinds of schools can ensure quality may, therefore, be misleading; a striving for improvement in a way that is not effective. As a general principle, proximal variables are more powerful than distal variables. It is what goes on in classrooms that creates quality in education not the nature of the building, the method of payment of the staff, or even the charisma of the Headteacher. The provision of high quality feedback directly to the classroom will, it is hypothesised, in a climate of investigation and an appeal to intrinsic motivation, lead to the highest quality system.