School Effectiveness and School Improvement
1999, Vol. 10, No. 2, pp. 217–232

0924-3453/99/1002-0217$15.00

SHORTER ARTICLES

Long-term Consequences of Curriculum Choices with Particular Reference to Mathematics and Science*

Carol Taylor Fitz-Gibbon
University of Durham

ABSTRACT

If *what* is taught is important but *how well* it is taught has only a trivial impact, then much of the work on 'school effectiveness' may be studying short term effects which quickly disappear from the system and have no long term consequences. If such were the case, then school effectiveness researchers would need to give greater consideration to what is studied, rather than simply how well it is studied. The influence of schools on curriculum choices may be more important than their influence on relative performance or "value added".

In the UK 'A' levels represent a useful point at which to look at the impact of curriculum choice since students typically have to choose to study only two or three subjects for the final 2 years of secondary school. The choices are made at the age of 16 with little evidence available regarding the long term consequences. This article presents an exploration of the consequences of taking or not taking 'A' level mathematics. Evidence was available from a 5 year follow-up study of students who took 'A' levels in 1988.

There were substantial differences between institutions in the extent to which students were attracted into mathematics, that is in the "Pulling Power" of mathematics departments. Focusing on students who appeared sufficiently able to have taken mathematics at A-level it was found that those who did and who were in high "Pulling Power" institutions, reported, 5 years later, a higher quality of life and a higher expectation for salaries than similarly able students who had been in institutions with low "Pulling Power" for mathematics and who had taken English at A-level.

In order to describe the education provided in a classroom, school, or a country, we need to state who is taught what, for how long, and how effectively. School effectiveness research has generally concentrated on the last item of information: how effectively are students taught? Thus

*This follow-up study was supported by the Economic and Social Research Council (grant number ESRC-R-000-23-3726-01 to Tymms & Fitz-Gibbon, 1993). The support of the Council is gratefully acknowledged as also are two helpful sets of comments from SESI reviewers. Responsibility for this report rests with the author.
An earlier version of this paper was presented at the annual meeting of the British Educational Research Association, September, 1994 in Oxford, England.

Correspondence: Carol Taylor Fitz-Gibbon, University of Durham, School of Education, Leazes Road, Durham DH1 1TA UK. Tel +44 191 374 3480. Fax +44 191 374 3506. E-mail C.T.Fitz-Gibbon@durham.ac.uk.

measures of pupil progress have been important and have come to be called "value added". Substantial differences have been found between different departments within schools and smaller differences have been found between schools (Aitkin & Longford, 1986; Fitz-Gibbon, 1985, 1991, 1992b; Gray, Jesson, & Jones, 1986; Nuttall, Goldstein, Prosser, & Rasbash, 1989; Smith & Tomlinson, 1989; Tymms 1992a, 1992b; Tymms & Fitz-Gibbon, 1991). It can be cogently argued, however, that *what* is taught may be of more consequence than how effectively it is taught. Which subjects were studied may have greater impact on the long-term knowledge of students, and may have more consequences for their subsequent life chances, than how well the subjects were studied. Preece (1983), studying achievement in science, suggested that the curriculum offered has far more impact than variations in the effectiveness of the instruction.

In addition to *personal* consequences the choices made of subjects to study will also, as they accumulate, have *national* consequences, such as was seen in the development of shortages in the supply of mathematics, science and foreign language teachers in the UK in the 1980s. Economists have considered the 'rate of return' following investment in education in general (e.g., Blaugh, 1965).

Currently, in the UK, considerable attention is paid to curriculum content through the work of the Examination Boards with their extensive procedures for consultation and curriculum revision. There is also, now, a National Curriculum which may be seen negatively, as limiting, or positively, as enabling, but which must certainly be seen as important if "content inclusion and emphasis" are as vital as Walker and Schaffarzick (1974) suggested. However, we also have "School Performance Tables" prepared by the Department for Education and Employment and published in the national press. These tables report a number of indicators, including the percentage of students who have five high-grade passes in examinations, but *without any differentiation between one subject and another*. The implication is that a pass is a pass, regardless of the content of the course. In England there are not, as yet, official School Performance Tables showing the balance of the curriculum even though such curriculum balance might be more important than whether or not the grades achieved show an extra point or two of "value added".

Effects of Differential Subject Difficulties and 'League Tables'

Should parents, teachers and careers officers 'encourage' particular choices of subjects post-16? If there is a continuing shortage of, for example, mathematicians, scientists and technologists, is it ethical for teachers and careers officers to encourage students to take up these subjects? One prob-

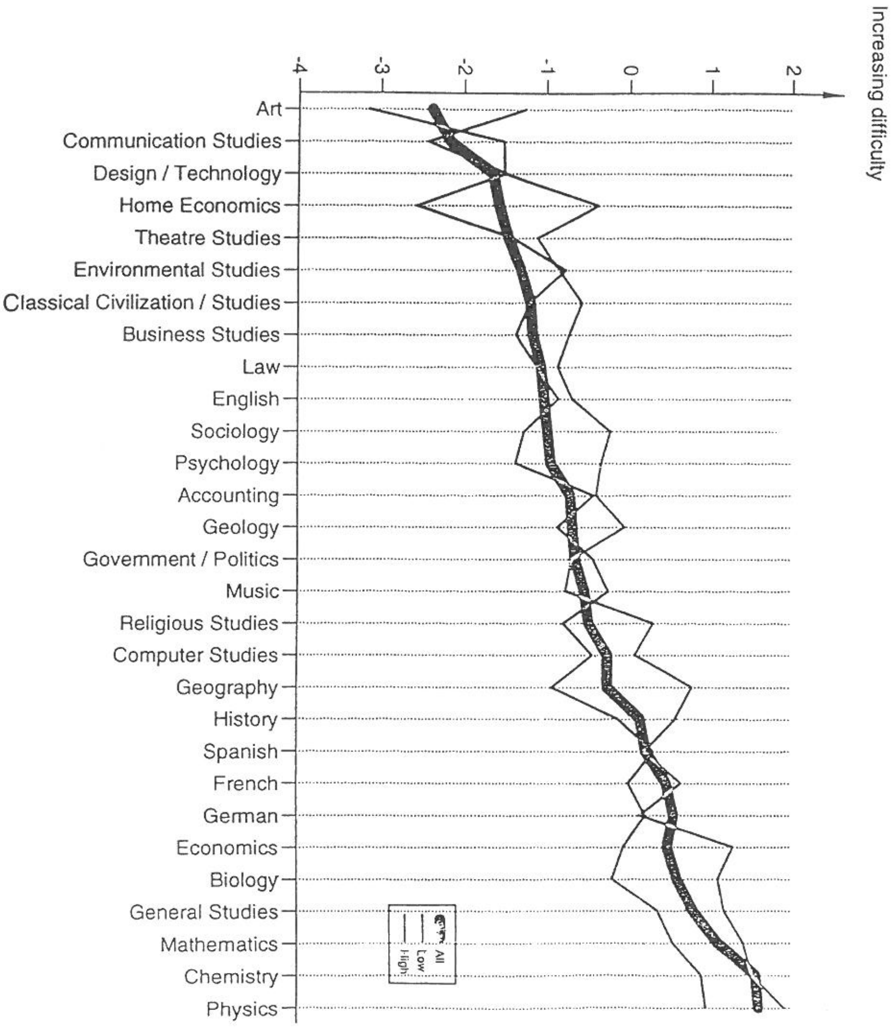

Fig. 1. Differences in difficulties of A-level subjects.

lem is that the grades students are likely to achieve in taking up mathematics, science or foreign languages are demonstrably lower than those that the same students would be likely to achieve doing other subjects (Fitz-Gibbon & Vincent, 1994).

The Correction Factor approach to measuring the difficulties of subjects, used in creating Figure 1, arises from work undertaken in Scotland by Alison Kelly and D.N. Lawley (Kelly, 1976). For each subject the Correction Factor indicates how much should be added to the obtained grade to produce the grade that would have been awarded had the examinations all been of equivalent difficulty. Thus Physics was the hardest, or most severely graded, subject and Art was the easiest, or least severely graded. (Further details can be found in Fitz-Gibbon & Vincent (1994)) The findings from this study were confirmed by analyses conducted by the Department for Education and Employment, using full national datasets, reported in the Dearing Report on the Education of 16 to 19 year olds, Dearing (1996).

There may be some compensation in the system in that the lower grades will, nevertheless, have the same opening power for places at universities: science departments accept students with lower grades on average than arts departments. However, if the effect of counselling students to take mathematics, science and foreign languages were to result in less positive outcomes subsequently for the students, either through failure at A-level, failure to obtain admission to chosen degree courses or failure to get jobs as good as those obtained by students studying arts subjects, then schools and careers officers might feel somewhat diffident in counselling students towards the more difficult subjects. Whether such worries would be justified is one of the concerns of this article.

In addition to the possible consequences of subject-choices for the *individual student* there are consequences for the school or college to be considered. Published School Performance Tables that make no differentiation between subjects, simply reporting indicators without reference to the subjects in which the grades were obtained, may well discourage schools and colleges from advising students to enrol in difficult subjects.

"Pulling Power"

In a report prepared in 1984 for 12 schools in the second year of the ALIS project[1] it was observed that the ratio of number of students taking mathematics to the number taking English was 1.52. The ratio nationally at that time was about 1.42. However, the ratio varied considerably, ranging from 0.69 in one school to 5.00 in another. That is to say, in some schools, there were fewer students in mathematics than in English classes, whilst in other schools there were as many as five students in mathematics for every one

1. Externally set and marked examinations are taken at the ages of 16 ("ordinary Level) and 18 (Advanced Level). The A-level Information System, ALIS, collects questionnaire data from schools and colleges, combines this with examination results and provides institutions with feedback on Value Added (residuals), students' attitudes and teaching and learning processes. The project started in 1983 as "Confidential, Measurement Based Self-Evaluation" and is run from the Curriculum, Evaluation and Management Centre, formerly at the University of Newcastle Upon Tyne and now at the University of Durham.

taking English. The term "Pulling Power" was coined to indicate the effect of departments that enrolled more than the usual proportion of students. The difficulty of interpreting the source of this variation was noted in the report:

> ...data alone cannot determine to what extent the ratios reflect the reputations established by the courses, the enthusiasms or canvassing activities of teachers or simply ...the value on the job market of subjects the candidate feels he or she has some chance of passing.
>
> (Fitz-Gibbon, 1984, p.10)

The differences in Pulling Power of mathematics departments represented very substantial differences between schools in the number of students that we might say were 'recruited' into mathematics. The use of the term 'recruited' is not meant to imply that any person actively campaigned to enrol students but rather that the students were in a school in which there was a higher probability of taking A-level mathematics than in other schools. The net effect was that they became part of a higher 'yield' of mathematically qualified students (Howson, 1987).

Because of the lower grades likely to be obtained in mathematics examinations, it seemed important, in studying Pulling Power, to recognise the danger students ran by choosing A-level mathematics and to select for study students who would stand a reasonable chance of passing A-level mathematics, that is to select for study those kinds of students who might reasonably be counselled towards mathematics-science subjects. The cut-off adopted to define "reasonably qualified" to study A-level mathematics was an average grade on the age-16 examinations which would place the student above the 25th percentile in the distribution of those who actually took mathematics A-level. For the data under consideration, from 1988, this cut-off point was 5.38 on a scale for the age-16 examinations (then called 'O-levels') in which A was assigned 7 points, B was assigned 6 points, et cetera.

Sources of Data

A dataset was established to follow up students from five north-east Local Education Authorities who had participated in the ALIS project in 1988 and had indicated a willingness to be contacted in the future. A 10-page questionnaire was completed by 47% (543) of those to whom a questionnaire was mailed. Sixty-one percent of those responding were female and this over representation of females may have affected the current analysis. An indicator of perceived 'Quality of Life' was assessed by a summated scale created by Tymms and based on the work of Marsh (Marsh, 1991; Tymms, 1995). The scale consisted of six items to which students responded on a 5-point scale ranging from "not true at all" to " very true" . The six items were:

(1) I like what I am doing
(2) I want to change my present position as soon as I can
(3) My colleagues are supportive
(4) I feel that I am treated well
(5) I would advise others to do what I am doing
(6) My present position is unsatisfactory.

The approach adopted was first to examine students who were in outlier institutions, extreme in terms of Pulling Power, and then to check the findings in the whole dataset.

INVESTIGATION OF EXTREME CASES

In a retrospective, passive observational study, there is considerable difficulty in distinguishing between students who might have been "recruited" into mathematical subjects in some way and others who could have been but were not. The strategy adopted was to identify institutions which were extremely high or extremely low in Pulling Power and study students in these institutions. Among these students there would be *some* who were 'pulled' towards, and some who were deterred from, the mathematics-science areas.

The 43 institutions from which there were responses to the follow-up questionnaire were rank-ordered on the simple ratio of A-level mathematics entries to English entries, that is on Pulling Power for mathematics as evidenced in the full dataset for 1988. Small institutions with fewer than 10 students taking mathematics or English were dropped from this part of the analyses. In order to obtain contrasting samples, institutions with a Pulling Power of 1 or less and institutions with a Pulling Power of 2.5 or more were selected. By also considering only candidates who were above the cut-off point mentioned above, that is were qualified by their O-level grade to take mathematics A-level, this method of selection sought out those who, with regard to taking A-level mathematics, could but didn't *and* were in low Pulling Power schools (the 'not pulled') and those who could and did *and* were in high Pulling Power schools (the 'pulled'). This selection for extremes — a selection which should have included some 'recruits' and 'failures to attract' — yielded 33 cases. Thirteen had taken

English in 3 institutions with low Pulling Power and 20 had taken mathematics in four institutions with high PP.

The Possibly 'Not-Pulled'

Of the 13 'qualified' students in *low* Pulling Power institutions who took English A-level, all were female and 3 had in fact taken A-level mathematics as well as English. Each one had failed. This suggested that institutions providing low yields in mathematics possibly did so by failing to attract qualified females into the subject, possibly because those who did take the subject did poorly. Table 1 shows residuals (based on multi-level modelling employing Average-GCSE score and sex as explanatory variables) for the mathematics-science departments in low and high Pulling Power institutions. It seems that the low Pulling Power mathematics departments were less effective than those in high Pulling Power institutions. In the sciences low Pulling Power institutions obtained modest results, more or less in line with expectations (residuals close to zero), but high Pulling Power institutions appeared to be particularly effective (large positive residuals). The ALIS data has generally shown no discrepancies among schools and colleges in their success with different gender groups (Tymms, 1992b) which leads to the hypothesis that the low Pulling Power institutions had generally poor results in mathematics, not just poor results for females.

The Possibly 'Pulled'

There were 20 students (13 males and 7 females) who took mathematics and were located in *high* Pulling Power institutions. Only one of these failed and only one had taken A-level English (she obtained a B in English, as compared with her D in mathematics.)

Outcomes for 'Pulled' and 'Not-Pulled' Students

Students taking mathematics in institutions with high Pulling Power were of similar prior achievement to the contrasting group, the not-pulled, unrecruited students (t=0.88, p=0.38). Thus 'Pulled' or 'recruited' students had shown higher achievement at age 16, but not reaching the usual levels of statistical significance despite an Effect Size of 0.31.

Having studied mathematics and other subjects at A-level, they then showed significantly higher A-level grades at age 18 (Effect Size 1.25). Their degree classifications were lower than those of non-pulled students, however, but not significantly so despite an Effect Size of –0.49. At the age of 23, respondents reported higher expected salaries (Effect Size 1.21) and a higher quality of life (Effect Size = 1.01) both statistically signifi-

Table 1. Low and High "Pulling Power" Institutions: Residuals and Intake Characteristics.

Low Pulling Power							
	Residuals				Intake Characteristics		
Inst. ID	Biology	Chemistry	Maths	Physics	Mean Prior Achievement	Mean SES	Type of Installation
307	–0.17	–0.08	–0.47	0.01	5.4	3.6	1
606	0.07	–0.15	–0.24	–0.22	5.5	4.4	3
610	–0.24	0.07	–0.36	–0.08	5.6	4.5	2

High Pulling Power							
	Residuals				Intake Characteristics		
Inst. ID	Biology	Chemistry	Maths	Physics	Mean Prior Achievement	Mean SES	Type of Installation
106	–0.16	–0.54	0.40	–0.08	5.6	4.0	3
107	1.18	0.33	0.91	0.18	5.6	3.6	3
304	0.32	0.89	0.21	0.25	5.5	4.5	1
702	0.51	0.76	–0.32	0.08	5.7	4.1	2

Notes. SES was measured on the Registrar General's scale for Head Of Household's Occupational Status and coded: 6= high down to 1= low. Pulling Power refers to the apparent success of mathematics departments in recruiting students, as compared with the English departments. 1 = 11–18 comprehensive school; 2 = 13–18 comprehensive school; 3 = Sixth Form college.

cant 'effects'. (The situation with regard to immediate salaries was confused because, at the time of the follow-up, some of the more academically successful would have still been on the low salaries of post-graduates working for higher degrees.)

Pupil Choice as Opposed to Institutional Pulling Power

The indicator of Pulling Power used so far has been an institution-level variable. We look now at the extent to which *an individual student* specialised in mathematics-science or had chosen easier subjects. The weights shown below were assigned to A-level subjects in order to create a measure of the extent to which difficult subjects had been chosen. The average weight for a student's choices was called the *Curriculum Choice* for that student.

Table 2. Outlier Analysis: Differences between Groups Containing Students who Were Not 'Pulled' (n=13) or Were, Possibly, 'Pulled' (n=20) into Mathematics.

Variable	n	Means	SD	t	two-tail p	Effect Size
Ach. at age 16						
Not pulled	13	5.92	0.39	0.89	0.38	0.32
pulled	20	6.04	0.38			
Ach. at age 18 (total points)						
Not pulled	13	7.31	4.58	3.44	0.002	1.25
pulled	19	12.95	4.45			
Degree classification						
Not pulled	9	3.78	1.00	3.28	0.233	–0.49
pulled	18	3.28	1.02			
Salary expected						
Not pulled	13	3.61	0.72	3.28	0.003	1.21
pulled	19	4.47	0.70			
Reported Quality of life at 23						
Not pulled	12	3.1	0.87	2.76	0.01	1.01
pulled	20	3.98	0.88			

Chemistry,	Physics and	Mathematics	5
Biology...	French...	German	4
Geography,	History,	Economics	3
English literature			2

The Curriculum Choice index represented predominantly the choice of science subjects since only about 12 to 15% chose foreign languages. Furthermore, a simple weighting of science against all others yielded correlations identical in the first decimal place with those from Curriculum Choice calculated as indicated above.

Relating Curriculum Choice to the salary expected in 5 years' time showed a linear and statistically significant relationship (r=.63 p<.0001, Fig. 2).

From the analysis of able students in outlier institutions with contrasting Pulling Power, it seemed that being 'pulled' into mathematics was associated with largely positive outcomes, with the exception of degree classification. We now turn from the analysis of outliers to the whole dataset.

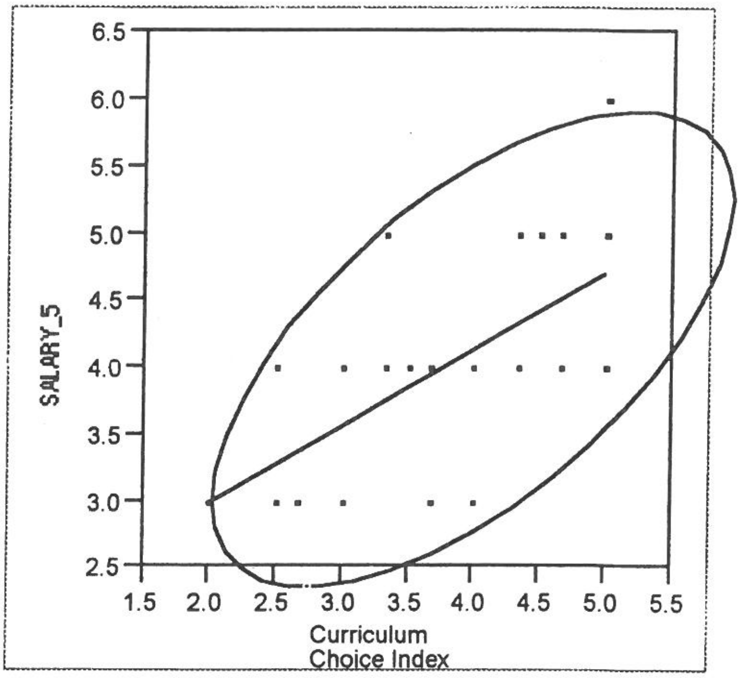

Fig. 2. Salary expected in the next 5 years related to the Curriculum Choice index in the A-levels. Choices were made by 32 students in low and high pulling power institutions (r=.63 p<.0001).

Note. Figure shows 90% ellipse and regression line
Salary levels were 1 = less than £6,000; 2 = £6,001 to £10,000; 3 = £10,001 to £15,000; 4 = £15,001 - £20,000; 5 = £20,000 to £30,000; 6= more than £30,000.

OUTCOMES IN THE WHOLE DATA SET

As would generally be expected, the patterns found by examining extreme groups were to some extent confirmed in the larger dataset although muted. Table 3 shows correlations based on 543 students.

From Table 3, we see that the Curriculum Choice index was quite strongly related to science performance at age 16, with a correlation of 0.55. This suggested that, if enrolment in the sciences is to be improved, interventions are probably needed earlier than at the age of 16 .

As in the analysis of extreme cases, curriculum choices weighted towards sciences did not seem to have endangered grade acquisition at the age of 18 (A-levels) nor the acquisition of a degree. In this larger dataset the correlation with degree classification was negligible. Bligh, Caves, & Settle (1980) reported that A-level grades correlated about 0.32 with degree classification and that by looking at within university-subject groupings (and thus removing differences in 'standards' or comparability of degrees across universities and subjects), the correlations ranged from .18 to .47. In Table 3 the correlation between degree classification and a sum of A-level points ("Achievement at age 18") of 0.29 was recorded. A-level grades do predict degree performance but the weakness of the correlation probably reflects the non-standardised nature of degree classifications. In other words degrees are no more "level" in difficulty than A-levels — probably less so.

There were significant correlations between the Curriculum Choice index and long-term consequences as evaluated by students responding to the questionnaires: a correlation of 0.26 with salary expected in 5 years and 0.10 with quality of life as perceived at the age of 23. the effects on outcomes were minimal and not significant for the indicator that is likely to be incorporated into School Performance Tables, namely the School Value Added measure or residual. The choices students make may be more influential than the effectiveness of the school they attend.

It might be thought that gender differences would account for some of these apparent effects but the patterns were highly similar for both male and female students (Table 3).

SUMMARY OF THE FINDINGS

The question posed was: since some schools already attract more students into mathematics-science subjects than do other schools, thus showing that the curriculum uptake *can* differ, and since there may be shortages in mathematics-science subjects, would it be ethical to urge other schools to attempt to increase the mathematics-science uptake? For informed discussion of this issue it would seem important to know whether being 'recruited' into mathematics-science (the difficult subjects) generally leads to failure, reduced life chances, and loss of earning power, or whether the effect is neutral or positive.

The findings from this follow-up study of more than 500 students suggested that

- Enrolment in mathematics-science subjects was strongly influenced by general academic level and prior achievement in the mathematics-science subjects but enrolments also differed considerably between institutions, some showing higher Pulling Power than others.

Table 3. Inter-correlations in the Follow-up Dataset (n=543).

	Ach at 16	SES	Asp. level	Prior science Ach.	Curriculum Choice	School Residual	Ach. at age 18	Degree Ach.	Degree standard	Salary expected in 5 yrs	Quality of life
At age 16											
Achievement		0.23	0.48	0.43	**0.23**	–0.00	0.63	0.02	0.20	0.19	0.04
SES	0.23		0.22	0.07	**0.06**	0.07	0.16	–0.00	0.03	0.08	–0.03
Aspiration level	0.48	0.22		0.28	**0.22**	0.05	0.49	0.05	0.20	0.18	0.05
Prior Ach. in Science	0.43	0.07	0.28		**0.55**	–0.01	0.28	–0.03	0.04	0.20	0.11
Curriculum Choice	**0.23**	**0.06**	**0.22**	**0.55**		**0.00**	**0.12**	**–0.01**	**–0.02**	**0.26**	**0.10**
At age 18											
School Residual	–0.00	0.07	0.05	–0.01	**0.00**		0.11	0.04	0.05	–0.04	0.05
Achievement	0.63	0.16	0.49	0.28	**0.12**	0.11		0.09	0.29	0.20	0.08
At age 21											
Degree achieved?	0.02	–0.00	0.05	–0.03	**–0.01**	0.04	0.09		0.32	–0.01	–0.00
Degree classification	0.20	0.03	0.20	0.04	**–0.02**	0.05	0.29	0.32		0.10	0.08
At age 23											
Salary expected in 5 yrs	0.19	0.08	0.18	0.20	**0.26**	–0.04	0.20	–0.01	0.10		0.24
Quality of Life	0.04	–0.03	0.05	0.11	**0.10**	0.05	0.08	–0.00	0.08	0.24	

Note.

Prior Achievement	Average O Level Grade at age 16
SES	Head Of Household's Occupational Status(6= high, 1= low)
Aspiration level	Likelihood of Staying in Education summated scale
Prior Ach. in Science	Sum of O level mathematics, physics and chemistry grades
Curriculum Choice	Weighted scale reflecting extent of choice of difficult subjects (mathematics, science and foreign languages. Predominantly affected by science choice)
School Residual	School residual (derived from multi-level model with prior ach. and Gender controlled)
Achievement at 18	Total Points at A-level taken at age 18.
Degree achieved?	Obtained a degree? Yes or no
Degree classification	Degree 'class', coded high is positive
Salary expected	Salary expected in 5 years' time, reported at the age of approximately 23
Quality of Life	Response on Quality of Life scale, answered at age of approximately 23

Table 4. Inter-correlations in the Follow-up Data Set, Boys and Girls Separately.

	Prior Achievement	Salary expected	School Residual	Curriculum Choice
Prior	1.00	0.23	–0.01	0.32
Achievement				
Salary expected	0.26	1.00	0.06	0.20
School Residual	–0.01	–0.00	1.00	0.02
Curriculum Choice	0.30	0.20	0.04	1.00

Note. Males below the diagonal; Females above the diagonal.

- Early achievement in mathematics-science areas was predictive of mathematics-science choices for A-level suggesting that any interventions or counselling might need to occur earlier than at the age of 16. Other studies have also suggested that this might be the case (Woolnough, 1995).
- Schools and colleges showed a wide range in the percentage of A-level entries falling into the mathematics-science category, from less than 45 to over 60%. The processes involved in generating these differences need investigation by both field work and survey analyses.
- For students who had been in schools or colleges with strong Pulling Power and had themselves included some mathematics-science subjects in their A-level choices (i.e., some of whom were possibly 're-cruits'), it was found that both their reported *quality of life* and *expected earnings* were higher than similar students not 'recruited' into mathematics-science A-level subjects. Whilst this must be seen as a tentative finding, confined to a few schools in the Northeast, over a limited period of time (5 years), it provides justification for *considering* whether more students might be counselled to take mathematics-science sub-

Table 5. Differences in Salary Variables for male and Female Students.

		Prior Achievement		Salary expected		School Residual		Curriculum Choice
SEXN	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*
Male	208	5.572	0.600	4.397	0.978	0.041	0.179	4.23
Female	317	5.577	0.666	3.878	0.835	0.043	0.175	3.62

jects at A-level, particularly in schools which are obtaining good progress (high residuals) in these areas.

- The Value Added indicator for a whole institution, a figure likely to be put into School Performance Tables to inform public choice of schools, was found to have little predictive validity for long-term consequences.

DISCUSSION AND CONCLUSIONS

It seems likely that curriculum choices have a continuing impact on life chances, employment, quality of life and other outcomes for students. Furthermore schools with strong Pulling Power towards mathematics appeared to be among the more effective in teaching mathematics and science subjects, a factor which may account for their Pulling Power.

Since one way to change a situation is to monitor it and feed the information back to the units responsible (Fitz-Gibbon, 1991,1992a, 1996; Tymms, 1991, 1992a, 1992b), one important outcome of this study is that the ALIS feedback to schools and colleges has been augmented to include analyses of each institution's curriculum balance in comparison with that in other schools and colleges. The balance between science, arts and mixed curriculum choices is now reported. Schools can see if they are unusual in the balance of the post-16 curriculum.

The present School Performance Tables that are published in the press in England, provide schools and colleges with a perverse incentive to keep students out of mathematics, sciences and foreign languages because these subjects will yield lower grades. Indeed, even when progress rather than end-point is reported the problem will remain as the Value Added is lower when sciences are chosen (Fitz-Gibbon & Vincent, 1994). The problem is ameliorated with regard to university entrance by admission to science degree courses being available for students with lower grades than those required for admission to arts degrees. A study which needs to be undertaken is the subsequent achievement in universities of students who study such subjects as economics, social science or medicine with or without having taken A-level mathematics. When such students encounter statistical parts of their disciplines, for example, are they at a severe disadvantage?

Since institutions such as schools may have substantial impacts, both on students and on society, it would seem important that long-term follow-up studies be conducted. The interpretations would be considerably more secure if based on controlled experiments but meanwhile, tentative conclusions have to be drawn from available data. From this dataset it can be suggested that there seems little reason to keep students away from difficult A-level subjects.

REFERENCES

Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society. Series A. 149,* 1–43.

Blaugh, M. (1965). The Rate of Return on Investment in Education. *The Manchester School, 33*(3), 205–251.

Bligh, D., Caves, R., & Settle, G. (1980). 'A' level scores and degree classifications as functions of university type and subject. In D. Billing (Ed.), *Indicators of Performance.* Guildford, Surrey: Society For Research Into Higher Education, 17–30.

Dearing, S.R. (1996). *Review of 16–19 Qualifications.* London: School Curriculum and Assessment Authority.

Fitz-Gibbon, C.T. (1984). *Combse 1984 A-level results: report to participating schools.* Newcastle Upon Tyne: the School of Education, the University.

Fitz-Gibbon, C.T. (1985). A-level results in comprehensive schools: The Combse project, year 1. *Oxford Review of Education 11*(1), 43–58.

Fitz-Gibbon, C.T. (1991). Multilevel modelling in an indicator system. In Raudenbush & Willms (Eds.), *Schools, Pupils and Classrooms: International Studies of Schooling from a Multilevel Perspective.* London and New York: Academic Press, 45–61.

Fitz-Gibbon, C.T. (1992a). The Design of Indicator Systems. *Research Papers in Education - Policy and Practice, 7*(3), 271–300.

Fitz-Gibbon, C.T. (1992b). School Effects at A-level: genesis of an information system. In D. Reynolds & P. Cuttance (Eds.), *School Effectiveness, Evaluation and Improvement* (pp. 96–120). London: Cassell.

Fitz-Gibbon, C.T. (1996). *Monitoring Education: Indicators, Quality and Effectiveness.* London, New York: Cassell.

Fitz-Gibbon, C.T., & Vincent L.S. (1994). *Candidates' performance in Science and mathematics at A-level.* London: School Curriculum and Assessment Authority.

Gray, J., Jesson, D., & Jones, B. (1986). The search for a fairer way of comparing schools' examination results. *Research Papers in Education 1*(2), 91–122.

Kelly, A. (1976). A study of the comparability of external examinations in different subjects. *Research in Education, 16,* 50–63.

Howson, G. (1987). *Challenge and Change.* Inaugural lecture given at the University of Southampton, January 1987. Faculty of Mathematical Studies.

Marsh, H.W. (1991). *Self description Questionnaire (SDQ) A theoretical and empirical basis for the measurement of multiple dimensions of late adolescent self concept: an interim test manual and a research monograph.* Macarthur, Australia: School of Education, University of Western Sydney.

Nuttall, D.L., Goldstein, M., Prosser, R., & Rasbash J. (1989). Differential School Effectiveness International. *Journal of Educational Research, 13,* 769–776

Preece, P.F.W. (1983). The qualitative principle of teaching. *Science Education, 67,* 69–73.

Smith, D.J., & Tomlinson, S. (1989). *The School Effect. A study of multi-racial comprehensives.* . London: Policy Studies Institute.

Tymms, P.B., & Fitz-Gibbon C.T. (1991). A comparison of exam boards: 'A' levels. *Oxford Review of Education, 17,* 17–31.

Tymms, P.B. (1991). Can Indicator Systems improve the effectiveness of Science and Mathematics Education? *Evaluation and Research in Education 4*(2), 61–73.

Tymms, P.B. (1992 a). The Relative Success of Post 16 Institutions in England (Including 'Assisted Places Schools'). *British Educational Research Journal, 18*(2), 175–192.

Tymms, P.B. (1992 b). Accountability – Can it be fair? *Oxford Review of Education, 19*(3), 291–299.

Tymms, P. B. (1995). The Long-term Impact of Schooling. *Evaluation and Research in Education, 9*(2), 99–108.

Walker, D.F., & Schaffarzick J. (1974). Comparing Curricular. *Review of Educational Research, 44*(1), 83–112.

Woolnough, B. (1995). School Effectiveness for different types of potential scientists and engineers. *Research in science and technology education, 13*(1), 53–66.